

DISSERTATION

Establishment of a proteomic workflow for biomarker research in blood cell samples

Ausgeführt zum Zwecke der Erlangung des akademischen Grades eines
Doktors der Naturwissenschaften
unter der Leitung von

Univ. Prof. Dr. Günter ALLMAIER
E164

Institut für Chemische Technologien und Analytik

Eingereicht an der Technischen Universität Wien
Fakultät für Technische Chemie

von

Mag. Wolfgang WINKLER
9725104

Jagdschlossgasse 54/9/7
1130 Wien

Wien, 10.12.2006

Zusammenfassung

Ziel der vorliegenden Dissertation war die Entwicklung eines Arbeitsablaufs für die Biomarker-Suche im Proteom von Blutzellen. Dazu wurden die Proteine der Zellen mittels zweidimensionaler Gelelektrophorese aufgetrennt, quantifiziert und jene Proteine, die zwischen den betrachteten Gruppen Unterschiede hinsichtlich ihres Expressionsverhaltens zeigten, mittels matrix-assisted laser desorption/ionisation (MALDI) Massenspektrometrie identifiziert.

Um eine gute und möglichst reproduzierbare Auftrennung der Proteine mittels zweidimensionaler Gelelektrophorese zu erzielen, ist eine Optimierung der Probenvorbereitung unerlässlich. Dazu wurde die generelle Tauglichkeit verschiedener Fällungsreagenzien und die Auswirkung ihrer Anwendung auf das Ergebnis der 2D-Elektrophorese untersucht, wobei sich die Verwendung von Trichloressigsäure als Fällungsmittel am besten bewährt hat.

Um eine Abschätzung der zu erwartenden Veränderungen am Proteom zu erhalten und die Proteinidentifizierung mittels Massenspektrometrie zu optimieren, wurde eine bereits bestehende Versuchsanordnung zur Analyse der Veränderungen am Proteom von U937-Zellen (eine monocytische Zelllinie) bei Glutamin-Mangel benutzt.

Die in den vorangegangenen Experimenten gewonnenen Erfahrungen wurden auf die Analyse primärer Zellen, Thrombocyten, übertragen. Um eine Biomarkersuche in Thrombocyten planen zu können, ist es notwendig Kenntnis der zwischen Proben von Individuen zu erwartenden Schwankungen hinsichtlich der Proteinexpression zu haben. Dazu wurden die Protein-Extrakte von Thrombocyten von 20 Freiwilligen analysiert und die Schwankung in den Konzentrationen der aufgetrennten Proteine innerhalb dieses Kollektivs bestimmt. Das Ausmaß dieser Schwankungen wurde dazu benutzt, die zumindest notwendige Probandenzahl abzuschätzen, mit der sich statistische Signifikanz für einen bestimmten Expressionsunterschied zwischen Gruppen von Individuen ergibt. Proteine, die besonders geringe Schwankungen aufweisen, wurden identifiziert und könnten in weiterer Folge für Normalisierungen bei anderen Methoden eingesetzt werden.

Aufbauend auf diese Erkenntnisse lässt sich eine Biomarker-Suche in Blutzellen, insbesondere Thrombocyten ohne weitere Vorversuche durchführen.

Acknowledgements

I would like to thank all who have contributed to the success of my work.

Special thanks go to:

My supervisors Prof. Dr. Rudolf Oehler and Prof. Dr. Günter Allmaier for their guidance, support and efforts to raise funds for projects.

The members of both working groups Mag. Roswitha Braunrath, Mag. Michael Diestinger, MSc Maja Eliassen, Mag. Tom Grunert, Mag. Jasmin Hirschmann, DI Christian Laschober, Dr. Martina Marchetti, Dr. Corina Mayrhofer, Dr. Roland Müller, Dr. Ernst Pittenauer, Mag. Senada Skeledzic, Dr. Gerald Stübiger, Dr. Martin Zehl and Dr. Maria Zellner for the good working atmosphere, valuable discussions and sharing of knowledge.

My parents for their support and understanding.

Table of Content

Abstract	6
Introduction.....	7
1 Compartments in Human Blood	7
1.1 Overview	7
1.2 Plasma	8
1.3 Lipoprotein Particles	10
1.4 Cells of the lymphoid lineage.....	11
1.4.1 B-Lymphocytes	11
1.4.2 T-Lymphocytes	12
1.4.3 Natural killer cells.....	12
1.5 Cells of the myeloid lineage.....	12
1.5.1 Granulocytes.....	12
1.5.2 Monocytes.....	13
1.5.3 Erythrocytes	14
1.5.4 Thrombocytes	15
1.6 References	18
2 Biomarkers and Biomarker Discovery	23
2.1 Overview	23
2.2 Evaluation of potential biomarkers.....	24
2.3 Guidelines for biomarker discovery studies	25
2.4 Influence of sample origin on biomarker discovery.....	26
2.5 References	27
3 Sources of Variation in the Biomarker Discovery Workflow	29
3.1 Workflow Description.....	29

3.2 Sources of Variation	30
3.3 References	32
Aims of this PhD Thesis	34
Published Papers and Manuscripts.....	35
1 Quantitative validation of different protein precipitation methods in proteome analysis of blood platelets.....	36
2 Adaptive cellular mechanisms in response to glutamine-starvation	46
3 Biological variation of platelet proteome in the elderly population and its implication for biomarker research.....	60
Conclusion	88
Lebenslauf.....	90

Abstract

The aim of this PhD-Thesis was the development of a workflow, suitable for biomarker search within the proteome of blood cells. Cellular proteins were separated by two-dimensional gel electrophoresis, quantified and those proteins, which showed different expression levels between the considered groups, were identified by matrix-assisted laser desorption/ionisation (MALDI) mass spectrometry.

To achieve good and reproducible separation of the proteins by two-dimensional gel electrophoresis, an optimisation of the sample preparation steps was necessary. Therefore, the applicability of different precipitation agents and their compatibility to 2D-electrophoresis was tested, whereas trichloroacetic acid has shown to be superior.

To get an estimation of the expression changes, that have to be expected and to optimise protein identification workflow based on by mass spectrometry, an already established experimental setup for the analysis of the changes in the proteome of U937 cells (monocytic cell line), that were challenged by glutamine starvation, was used.

The experiences that have been gained during the preliminary experiments were applied to the analysis of primary cells, thrombocytes. To design a search for biomarkers in thrombocytes, it is necessary to assess the amount of variability in protein expression levels between samples of different individuals. Therefore, the thrombocyte protein extracts, obtained from 20 volunteers, were analysed and the variation in protein expression levels within this group was determined. This degree of variation was used to estimate the minimum number of samples, that is needed to obtain statistical significance for a given difference in expression levels when comparing two groups of individuals. Some proteins showed very low variation and were identified by mass spectrometry. These proteins and can serve for normalisation purposes in different methods.

Taking into account these results we are now able to define the technical procedure to perform a biomarker search in blood cells without further experiments.

Introduction

1 Compartments in Human Blood

1.1 Overview

Blood is classified as connective tissue and circulates through the veins of all vertebrates by the pumping action of the heart. Blood consists of several types of highly specialised cell types suspended in a liquid intercellular fluid, the blood plasma.

The main functions of blood are:

- Supply of tissues with oxygen and the removal of carbon dioxide
- Supply of tissues with nutrients, the transport of metabolites between different tissue types or organs and the removal of metabolic waste products
- Transport and distribution of substances with regulatory and signalling function
- Support of immune defence by circulating white blood cells and antibodies
- Support the maintenance of blood vessel integrity by circulating thrombocytes and coagulation factors
- Regulation and maintenance of water and electrolyte balance
- Regulation of pH-value and body core temperature

An adult human body typically contains a blood volume of 5-6 litres, it shows a density of about 1.06 g/ml and a pH value of approximately 7.4 [1]. Typically 55% of the blood volume is made up by plasma which is an aqueous solution of proteins and several types of small molecules like peptides, salts, sugars or amino acids.

The remaining 45% of blood volume is made up by different types of cells and lipoprotein particles, the total cell count is about $4.4 \times 10^9 - 6.3 \times 10^9$ cells/ml. With $4.2 \times 10^9 - 5.9 \times 10^9$ cells/ml the erythrocytes are the most frequent blood cell type,

followed by thrombocytes ($1.7 \times 10^8 - 4 \times 10^8$ cells/ml) and leukocytes ($5 \times 10^6 - 1 \times 10^7$ cells/ml) [2].

The biogenesis of blood cells takes place in the bone marrow. It is estimated that around 2×10^{11} erythrocytes and 10^{10} leukocytes have to be produced per day in adult humans to keep up the blood cell levels [3]. All types of blood cells descend from the pluripotential hematopoietic stem cells via several steps of differentiation and maturation, mediated by hematopoietic growth factors [3-5]. The pluripotential hematopoietic stem cells give rise to the common progenitor cells of the lymphoid and myeloid lineage. The common lymphoid progenitor cells give rise to B-lymphocytes, T-lymphocytes and natural killer cells. The common myeloid progenitor cells give rise to the granulocytic/monocytic progenitor cells which differentiate into granulocytes and monocytes and the megakaryotic/erythroid progenitor cells which differentiate into the platelet producing megakaryocytes and erythrocytes.

According to their different functions, blood cells have very different lifespans ranging from one or two days (granulocytes) to months (erythrocytes) or years (some lymphocytes) [2].

Blood is the primary source for clinical diagnostics, because blood samples are easy and repeatedly to obtain in more than sufficient amounts. Due to its functions, blood is directly connected to every tissue in the body and therefore in principle carries information about each tissues state [6, 7]. In addition to diagnostic use, blood is used for therapeutic purposes, either as whole blood transfusion or as a drug obtained by processing blood or one of its components.

1.2 Plasma

Blood plasma is obtained as a yellowish liquid after centrifugation of blood treated with anticoagulants.

It is an aqueous solution of:

- Salts: Mainly sodium, chloride, bicarbonate, phosphates; are responsible for establishing the osmotic pressure and contribute to pH buffering (bicarbonate and phosphates)

- Small molecules: Metabolites like glucose, organic acids, amino acids, metabolic waste products like urea and messengers like hormones
- Proteins and peptides which contribute to immune defence, coagulation, or are responsible for specific metabolite transport. Additionally the plasma proteins are responsible for the colloidal osmotic pressure which controls the volume balance between blood and interstitial fluid.

The group of proteins and peptides (about 70 g/l plasma) shows the by far highest diversity, by means of the number of protein and peptide species as well as the dynamic range of their abundance (spans at least 12 orders of magnitude) [8]. Unfortunately, about 99% of the total protein content is made up of the 22 most abundant proteins [9]: albumin (transport protein; 50% of the total protein content), total IgG (immune defence), transferrin (iron transport), fibrinogen (coagulation), total IgA (immune defence), α -2-Macroglobulin (protease inhibition), total IgM (immune defence), α -1-antitrypsin (protease inhibition), complement C3 (immune defence), haptoglobin (haemoglobin degradation), apolipoprotein A-1 and B (lipid metabolism), α -1-acid glycoprotein (immune defence), lipoprotein(a) (lipid metabolism), Factor H (immune defence), Ceruloplasmin (iron metabolism), complement C4 (immune defence), complement Factor B (immune defence), Transthyretin (transport protein), complement C9 and C1q and C8 (immune defence). The presence of a few extreme high abundant proteins (mg/ml-range) next to low abundant potential biomarker proteins (pg/ml and below) is a major challenge for current separation and analysis methods and requires the development of strategies either to deplete these proteins or analyze the low abundant proteins in presence of the high abundant ones [10]. But neither unspecific depletion methods like precipitation nor specific depletion by immunoaffinity can avoid the co-removal of other proteins which interact specifically or unspecifically with the proteins intended to remove or the stationary phase [10]. Using a strategy of immunoaffinity fractionation followed by two-dimensional chromatographic separation and identification by mass spectrometry, more than 4500 proteins could be identified [11]. About 55% of the identified proteins were classified as cellular proteins (intracellular, secreted cellular and membrane associated proteins) and are caused by cellular leakage.

Recent biomarker studies using screening methods like SELDI [12, 13] have drawn attention to the peptide and low molecular protein fraction of human plasma. Again,

like observed for the high molecular weight fraction, compounds unexpected to be present in plasma, like peptides from structural and nuclear proteins, were identified next to anticipated peptides like fragments from plasma proteins and biologically active peptides and peptide precursors [9, 14].

1.3 Lipoprotein Particles

Lipoprotein particles are spheres with a hydrophilic surface consisting of a phospholipid layer stabilised by different types of apolipoproteins and a hydrophobic core consisting of triacylglycerides, cholesterol, cholesterol esters. The lipoprotein particles are classified by their density (Chylomikrons, very low density lipoprotein (VLDL), low density lipoproteins (LDL), intermediate density lipoproteins (IDL) and high density lipoproteins (HDL). The main function of all types of lipoprotein particles is transport and short term storage of lipids and cholesterol as well as the supply of tissues with triacylglycerides and cholesterol.

Fat out of food is resorbed by mucosa cells of the small intestine, and secreted as chylomikrons into the lymphatic vessels. Chylomikrons are the largest lipoprotein particles ($\varnothing=100-1000$ nm), consist mainly of triacylglycerides (85%) and their main apolipoprotein is Apo B-48. They enter the blood stream at the thoratic duct's connection with the left subclavian vein. During their circulation time (about 10 minutes) chylomikrons receive apolipoproteins A and C from HDL and their triacylglyceride content is cleaved by lipoprotein lipases which are located at the surface of blood vessel endothelial cells and adipose tissue cells. The cleavage products are subsequently resorbed by tissues and the chylomikron remnants are resorbed by the liver and degraded. The remaining triacylglycerides are again released by the liver incorporated in VLDL ($\varnothing=50$ nm) containing triacylglycerides, cholesterol esters and apolipoproteins B100, C-II and E. Their triacylglyceride content is again reduced by lipoprotein lipase cleavage and resorbtion by tissues and VLDL are transformed into IDL which are either resorbed by the liver and re-released as VLDL or their triacylglyceride content is further reduced by hepatic tissues transforming IDL into LDL. LDL ($\varnothing=20$ nm) circulate for about 5 days, consist of cholesterol esters, triacylglycerides, and apolipoprotein B-100 and provide cholesterol to tissues until they are degraded by the liver. Unlike all other lipoprotein particles which provide cholesterol and lipids to tissues, HDL are released from the

liver to collect and remove excess cholesterol from tissues and ensure its degradation. HDL also interacts and exchanges apolipoproteins with other lipoprotein particles.

Lipoprotein particles and their involvement in arteriosclerosis is subject of ongoing investigation [15]. Recent proteomic studies extended the knowledge about involved proteins and the extent of their modification [16-19].

1.4 Cells of the lymphoid lineage

The lymphocytes are the cells of the lymphoid lineage. They can be divided into three major types: B-lymphocytes, T-lymphocytes and natural killer cells. B- and T-lymphocytes belong to adaptive immune system and natural killer cells belong to the innate immune system.

Recent proteomic studies created proteome maps of T-helper cell types [20], general signalling processes in lymphocytes [21] and the proteins and their reorganisation in lipid raft proteome during T-cell activation [22, 23].

1.4.1 B-Lymphocytes

The complete biogenesis of B-lymphocytes takes place in the bone marrow: The lymphoid progenitor cell give rise to the pro-lymphocytes which further differentiate into pre-lymphocytes. These early stages is accompanied by high mitotic activity. Pre-lymphocytes differentiate into the small (immature) lymphocytes which further differentiate into mature B-lymphocytes.

Mature B-lymphocytes carry IgM and IgD on their surface which are directed against one single antigen. B-lymphocytes become activated if this antigen is recognized by the surface bound receptors, and the cell additionally receives activation signals from helper-T-cells. Fully activated B-lymphocytes differentiate either into plasma B cells, which are responsible for antibody production, or into memory B cells which are responsible for long-term immunity.

1.4.2 T-Lymphocytes

The biogenesis of T-lymphocytes is similar to the biogenesis of B-lymphocytes, but takes place in the thymus.

In difference to B-lymphocytes, T-lymphocytes don't recognise free antigens, but only antigens that are presented by antigen presenting cells.

T-lymphocytes are further classified into helper T cells (T_H1 cells activate macrophages, T_H2 cells activate B-lymphocytes), T-regulator cells (responsible for immunological tolerance) and cytotoxic T-cells (induce apoptosis in cells with viral infections).

1.4.3 Natural killer cells

Natural killer cells belong to the lymphoid lineage and are produced and mature in the bone marrow. They belong to the innate immune system and specifically recognise cells with viral infection or tumour cells. Natural killer cells induce apoptosis in these cells.

1.5 Cells of the myeloid lineage

1.5.1 Granulocytes

The types of granulocytes are distinguished by their different appearance in light microscopy after staining.

Granulocyte biogenesis takes place in the bone marrow and is similar for all different types of granulocytes. The common granulocyte/monocyte progenitor cells give rise to myeloblasts, which proliferate to form promyelocytes which give rise to myelocytes. Here cell proliferation and the expression of proteins for primary granulae stop, and the biogenesis of secondary and tertiary granulae as well as the synthesis of their protein content begins. Myelocytes develop into metamyelocytes which differentiate into band cells and finally into mature granulocytes [24, 25]. As the maturation proceeds, cell size and nuclear shape undergo distinct changes.

Several proteomic studies investigated different aspects like granulae composition [26] and membrane skeleton [27] of neutrophils and the activation processes of eosinophils [28].

Neutrophil granulocytes

Neutrophil granulocytes are phagocytes who circulate until they either are activated by an inflammation or commit apoptosis after about 6 hours. Activated neutrophils can leave the blood vessels and migrate into tissues to phagocyte pathogens. Neutrophils contain azurophil and specific granulae which merge with the phagosome to support the degradation of pathogens. Additionally the content of specific and gelatinase granulae is released to the extracellular environment [24].

Eosinophil granulocytes

Eosinophil granulocytes are recruited from blood stream into late phase reaction sites or sites of helminthic infection via a combination of adhesion molecule interaction and chemokines [29]. Their granulae contain lysosomal hydrolases as well as eosinophil-specific peroxidase and proteins with toxic properties to helminths, bacteria and normal tissue cells.

Basophil granulocytes

Basophil granulocytes are recruited from blood stream into inflammatory sites, usually together with eosinophils. Basophils are activated by antigen-bound IgE and release their granulae content (Histamine and enzymes which degrade microbial structures) and additionally release lipid mediators and cytokines produced upon activation [29].

1.5.2 Monocytes

Monocytes are phagocytes, which are derived from granulocytic/monocytic progenitor cells via the intermediary state of monoblasts in the bone marrow. Monocytes have a characteristically shaped nucleus and contain many intracytoplasmatic lysosomes and often azurophilic granulae. *In vitro*, monocytes show strong adhesion to glass and plastic surfaces and they actively phagocyte organisms. Receptors on the monocyte surface bind to certain carbohydrate structures on the microbial cell walls or to IgG and proteins of the complement system which have coated the microbe.

After their release into the blood stream, monocytes circulate for 1 – 3 days until they migrate into tissues and develop into different types of macrophages, depending on the tissue type. Neither monocytes nor macrophages are terminally differentiated and retained the ability to perform cell division.

Besides other tissues and organs like liver (called Kupffer cells) and spleen (called sinusoidal lining cells), where macrophages remove waste products by phagocytosis, they are placed at all sites where microbes can enter the host. Like neutrophil granulocytes, macrophages recognise microbes by interaction with surface proteins triggering the phagocytosis of the microbe.

In contrast to granulocytes, monocytes and monocyte derived cells are antigen presenting cells. They have the ability to present a fragment of the phagocytosed microbe on their surface. This leads to activation of T-cells and finally also of B-cells. Thus, monocytes represent the interconnection of the innate immune system with the adaptive immune system.

To study monocyte/macrophage differentiation in vitro, the cell line U937 has been established as a model system. This cell line is derived from a diffuse histiocytic lymphoma and has many monocytic characteristics [30].

A couple of proteomic studies investigating monocytes, macrophages or U937 cells have been done so far: Secretome and cytoplasmatic proteome of macrophages [31], proteome comparison macrophages – monocytes [32], analysis of the response of U937 to oxidative stress [33] and glutamine starvation [34].

1.5.3 Erythrocytes

Mammalian erythrocytes are biconcave disks with a diameter of approximately 7.5 μm . They are the most frequent blood cells (4.2×10^9 - 5.9×10^9 cells/ml) and are responsible for the supply of tissues with oxygen and the removal of carbon dioxide. The oxygen uptake in the lung and its release in tissues is accomplished by the protein haemoglobin of which erythrocytes contain up to 3×10^8 copies per cell. The disk-like shape of erythrocytes is advantageous for oxygen exchange and their cytoskeleton makes changes of their shape possible to ease the passage through capillaries. Erythrocytes carry the antigens (terminal glyco-structures linked to ceramide anchor) determining the blood type of each individual on their surface. Most

important for blood transfusion are the antigens described by the AB0 and Rhesus systems.

Erythrocytes belong to the myeloid lineage and their biogenesis leaves the common pathway with the differentiation of the common megakaryocytic/erythroid progenitor cells into erythroid burst forming units (BFU-E), which further differentiate into erythroid colony forming units (CFU-E). Erythropoietin (EPO) amongst several other growth factors induces the differentiation of CFU-E into the proerythroblast which is the first cell type of the erythropoiesis pathway [35]. Proerythroblasts proliferate under control of EPO into basophilic erythroblasts which heavily synthesise haemoglobin and mature into the polychromatic erythroblasts which are the last developmental stage that is capable of performing cell division. The orthochromatic erythroblast that emerges this last cell division condenses its nucleus which is subsequently expelled along with mitochondria, endoplasmic reticulum and ribosomes. These anucleated reticulocytes are still capable of haemoglobin synthesis for some time and are released into the blood stream where they develop into mature erythrocytes.

The average lifetime of an erythrocyte in circulation is about 120 days until they are degraded in spleen and liver.

For decades now, erythrocytes and their proteins have been object of intense research. Recent studies used proteomics techniques to study the erythrocyte membrane [36, 37], the cytoplasmatic proteome [37] as well as aspects of erythropoiesis [38] and malaria [39].

1.5.4 Thrombocytes

Thrombocytes are the smallest cellular blood components, they are disk-like shaped ($3.6 \times 0.7 \mu\text{m}$) and appear in human blood in concentrations of 1.7×10^8 – 4×10^8 cells/ml. Their primary function is to maintain the vascular integrity, but they are also involved in processes of the innate immune defence, inflammatory response and wound healing [40, 41]

Thrombocyte biogenesis takes place in the bone marrow, where they are released from megakaryocytes as cytoplasmatic fragments. Megakaryocytes belong to the myeloid lineage and descend from the common megakaryocytic/erythroid progenitor

cells. These progenitor cells differentiate into megakaryocytic burst forming units (BFU-MK) which give rise to megakaryocytic colony forming units (CFU-MK). CFU-MK give rise to the megakaryoblasts which lose their ability to perform cell division but retain the ability to perform DNA replication resulting in ploidy levels of up to 128N [42, 43]. During transformation into megakaryocytes, megakaryoblasts accumulate platelet specific proteins either by synthesis or endocytosis from blood plasma, as well as organelles, granulae and the demarcation membrane system [44, 45]. In presence of growth factors especially thrombopoietin (TPO) megakaryocytes change their shape by forming pseudopodial extensions (proplatelets) of 100-500 μm length and a diameter of 2-4 μm which undergo repeated cycles of extension, bending and branching until the megakaryocyte's whole cytoplasmic content is transformed into a network of interconnected proplatelets [46]. Organelles and granulae are actively transported along microtubules through the proplatelets into the nascent platelets, formed on the very end of each proplatelets, where they are captured [47]. Finally, single platelets or short chains of proplatelets fragments are released from the proplatelet network into circulation by rapid retraction. Platelets circulate for 10-12 days until they are degraded by spleen and liver.

In the case of a blood vessel injury, subendothelial matrix is exposed to the platelets which specifically interact with the subendothelial matrix via their membrane glycoproteins (GP) leading to an adhesion of platelets at the site of injury. This is accomplished either by direct binding of the platelet membrane glycoprotein GPIa/IIa to the subendothelial protein collagen or by the binding of platelet membrane complex GPIb/V/IX to the plasma protein von Willebrand factor (vWF) which has previously bound to collagen [40]. Adhesion triggers platelet activation which causes a change in shape and the secretion of substances contained in α -granulae (contain adhesion molecules, chemokines growth factors, proteins of pro-coagulant and fibrinolytic system, immunological and other proteins) and dense granulae (contain ions, nucleotides, membrane proteins and transmitters) [48]. The secreted substances lead to activation and attraction of further platelets resulting in platelet aggregation and the formation of a haemostatic plug where activated platelets are linked by fibrinogen bridges bound to their membrane glycoproteins GPIIb/IIIa. The membrane of activated platelets provides the necessary enzymes and cofactors to run the coagulation cascade yielding a burst of thrombin: The Xase-complex

(factor IX together with the platelet membrane bound factor VIII) activates factor X which forms the prothrombinase complex (factor X together with platelet membrane bound factor V). The prothrombinase complex activates thrombin which transforms fibrinogen into fibrin, directly on the site of vessel injury. Fibrin monomers aggregate to fibrils which stabilise the platelet clot and are subsequently cross-linked by the enzymatic activity of factor XIII.

To ensure a local limitation of coagulation, avoid an excessive coagulation or an initiation of coagulation without adequate cause, several substances have anticoagulatory effect. Prostacyclin (released by endothelial cells) and its antagonist thromboxane (released by platelets upon activation) inhibit and promote the platelet aggregation, respectively. Antithrombin especially when interacting with heparan sulfate is the most important inhibitor of coagulation, it inactivates factor X and thrombin by proteolytic cleavage. Protein C is activated by thrombin and forms a complex with protein S. This complex is attached to the membranes of activated platelets via Calcium ions and inactivates factor V and factor VIII by proteolytic cleavage.

Once, the damaged blood vessel wall has been regenerated after injury, the clot has to be removed to re-establish normal blood flow. The main component of the fibrinolytic system is plasmin which degrades the fibrin fibrils by proteolytic cleavage. Plasmin activation and inactivation is controlled by plasminogen activators (tissue-type plasminogen activator, urokinase), activator-inhibitors (PAI-1) and plasmin inhibitors (α -antiplasmin) [49, 50].

Due to their essential function in the process of haemostasis, platelets have been in focus of research for many years. Epidemiological studies show a dramatic increase in the rates of thrombotic events like myocardial infraction and stroke in the population with advancing age [51]. These findings obviously are the result of an age-related disorder in the haemostatic system. Alterations in protein expression levels which suggest an increasing imbalance between pro- and anticoagulant systems, as well as alterations in the fibrinolytic system have been described [52]. Besides their obvious involvement in many disorders of the haemostatic system, platelets can also serve as system to study even neurodegenerative diseases like Alzheimer's disease [53].

Being anucleate cytoplasmatic fragments, genomic methods can't unleash their full power to investigate thrombocytes. During the process of biogenesis, thrombocytes have received a set of mRNA from the megakaryocytes and are able to perform protein biosynthesis [54]. Therefore platelets can not respond to external stimuli by induction of new proteins, but they have limited potential to modify the translated mRNA by signal dependent splicing [55]. Recent studies have revealed the presence of 2000 – 3000 transcripts in platelets [54, 56].

Several proteomic studies applying various techniques have focussed on different aspects of platelet research. Gel based or gel free peptide or protein separation followed by mass spectrometry based protein identification was used to analyse the static platelet proteome [57-60], the platelet secretome [61], platelet membrane proteins [62] and platelet phosphoproteins [63, 64].

1.6 References

1. Horn, F., et al., *Biochemie des Menschen*. 3rd Edition ed. 2005, Stuttgart: Georg Thieme Verlag.
2. Klinke, R., H.-C. Pape, and S. Silbernagl, eds. *Physiologie*. 5th Edition ed. 2005, Georg Thieme Verlag: Stuttgart.
3. Bellantuono, I., *Haemopoietic stem cells*. *Int J Biochem Cell Biol*, 2004. **36**(4): p. 607-20.
4. Passegue, E., et al., *Normal and leukemic hematopoiesis: are leukemias a stem cell disorder or a reacquisition of stem cell characteristics?* *Proc Natl Acad Sci U S A*, 2003. **100 Suppl 1**: p. 11842-9.
5. Akashi, K., et al., *A clonogenic common myeloid progenitor that gives rise to all myeloid lineages*. *Nature*, 2000. **404**(6774): p. 193-7.
6. Jacobs, J.M., et al., *Utilizing human blood plasma for proteomic biomarker discovery*. *J Proteome Res*, 2005. **4**(4): p. 1073-85.
7. Liotta, L.A., M. Ferrari, and E. Petricoin, *Clinical proteomics: written in blood*. *Nature*, 2003. **425**(6961): p. 905.
8. Anderson, N.L. and N.G. Anderson, *The human plasma proteome: history, character, and diagnostic prospects*. *Mol Cell Proteomics*, 2002. **1**(11): p. 845-67.
9. Tirumalai, R.S., et al., *Characterization of the low molecular weight human serum proteome*. *Mol Cell Proteomics*, 2003. **2**(10): p. 1096-103.

10. Tam, S.W., J. Pirro, and D. Hinerfeld, *Depletion and fractionation technologies in plasma proteomic analysis*. Expert Rev Proteomics, 2004. **1**(4): p. 411-20.
11. Shen, Y., et al., *Characterization of the human blood plasma proteome*. Proteomics, 2005. **5**(15): p. 4034-45.
12. Hutchens, W.T. and T.-T. Yip, *New desorption strategies for the mass spectrometric analysis of macromolecules*. Rapid Communications in Mass Spectrometry, 1993. **7**(7): p. 576-580.
13. Adam, B.L., et al., *Serum protein fingerprinting coupled with a pattern-matching algorithm distinguishes prostate cancer from benign prostate hyperplasia and healthy men*. Cancer Res, 2002. **62**(13): p. 3609-14.
14. Richter, R., et al., *Composition of the peptide fraction in human blood plasma: database of circulating human peptides*. J Chromatogr B Biomed Sci Appl, 1999. **726**(1-2): p. 25-35.
15. Steinberg, D., *Atherogenesis in perspective: hypercholesterolemia and inflammation as partners in crime*. Nat Med, 2002. **8**(11): p. 1211-7.
16. Farwig, Z.N., A.V. Campbell, and R.D. Macfarlane, *Analysis of high-density lipoprotein apolipoproteins recovered from specific immobilized pH gradient gel pl domains by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry*. Anal Chem, 2003. **75**(15): p. 3823-30.
17. Ogorzalek Loo, R.R., et al., *Virtual two-dimensional gel electrophoresis of high-density lipoproteins*. Electrophoresis, 2004. **25**(14): p. 2384-91.
18. Karlsson, H., et al., *Lipoproteomics I: mapping of proteins in low-density lipoprotein using two-dimensional gel electrophoresis and mass spectrometry*. Proteomics, 2005. **5**(2): p. 551-65.
19. James, R.W., et al., *Protein heterogeneity of lipoprotein particles containing apolipoprotein A-I without apolipoprotein A-II and apolipoprotein A-I with apolipoprotein A-II isolated from human plasma*. J Lipid Res, 1988. **29**(12): p. 1557-71.
20. Nyman, T.A., et al., *A proteome database of human primary T helper cells*. Electrophoresis, 2001. **22**(20): p. 4375-82.
21. Wollscheid, B., J.D. Watts, and R. Aebersold, *Proteomics/genomics and signaling in lymphocytes*. Curr Opin Immunol, 2004. **16**(3): p. 337-44.
22. Wollscheid, B., et al., *Lipid raft proteins and their identification in T lymphocytes*. Subcell Biochem, 2004. **37**: p. 121-52.
23. Bini, L., et al., *Extensive temporally regulated reorganization of the lipid raft proteome following T-cell antigen receptor triggering*. Biochem J, 2003. **369**(Pt 2): p. 301-9.

24. Theilgaard-Monch, K., et al., *The transcriptional program of terminal granulocytic differentiation*. Blood, 2005. **105**(4): p. 1785-96.
25. Van Merris, V., E. Meyer, and C. Burvenich, *Functional maturation during bovine granulopoiesis*. J Dairy Sci, 2002. **85**(11): p. 2859-68.
26. Lominadze, G., et al., *Proteomic analysis of human neutrophil granules*. Mol Cell Proteomics, 2005. **4**(10): p. 1503-21.
27. Nebl, T., et al., *Proteomic analysis of a detergent-resistant membrane skeleton from neutrophil plasma membranes*. J Biol Chem, 2002. **277**(45): p. 43399-409.
28. Levi-Schaffer, F., et al., *Proteomic analysis of human eosinophil activation mediated by mast cells, granulocyte macrophage colony stimulating factor and tumor necrosis factor alpha*. Proteomics, 2002. **2**(11): p. 1616-26.
29. Abbas, A., A. Lichtman, and J. Pober, *Cellular and Molecular Immunology*. 4th ed. 2000.
30. Zellner, M., et al., *Glutamine starvation of monocytes inhibits the ubiquitin-proteasome proteolytic pathway*. Biochim Biophys Acta, 2003. **1638**(2): p. 138-48.
31. Dupont, A., et al., *Two-dimensional maps and databases of the human macrophage proteome and secretome*. Proteomics, 2004. **4**(6): p. 1761-78.
32. Jin, M., et al., *Proteome comparison of alveolar macrophages with monocytes reveals distinct protein characteristics*. Am J Respir Cell Mol Biol, 2004. **31**(3): p. 322-9.
33. Seong, J.K., et al., *Proteomic analysis of the cellular proteins induced by adaptive concentrations of hydrogen peroxide in human U937 cells*. Exp Mol Med, 2002. **34**(5): p. 374-8.
34. Eliassen, M.M., et al., *Adaptive cellular mechanisms in response to glutamine-starvation*. Front Biosci, 2006. **11**: p. 3199-211.
35. Fisher, J.W., *Erythropoietin: physiology and pharmacology update*. Exp Biol Med (Maywood), 2003. **228**(1): p. 1-14.
36. Low, T.Y., T.K. Seow, and M.C. Chung, *Separation of human erythrocyte membrane associated proteins with one-dimensional and two-dimensional gel electrophoresis followed by identification with matrix-assisted laser desorption/ionization-time of flight mass spectrometry*. Proteomics, 2002. **2**(9): p. 1229-39.
37. Kakhniashvili, D.G., L.A. Bulla, Jr., and S.R. Goodman, *The human erythrocyte proteome: analysis by ion trap mass spectrometry*. Mol Cell Proteomics, 2004. **3**(5): p. 501-9.

38. Brand, M., et al., *Dynamic changes in transcription factor complexes during erythroid differentiation revealed by quantitative proteomics*. Nat Struct Mol Biol, 2004. **11**(1): p. 73-80.
39. Florens, L., et al., *A proteomic view of the Plasmodium falciparum life cycle*. Nature, 2002. **419**(6906): p. 520-6.
40. Jurk, K. and B.E. Kehrel, *Platelets: physiology and biochemistry*. Semin Thromb Hemost, 2005. **31**(4): p. 381-92.
41. Weyrich, A.S. and G.A. Zimmerman, *Platelets: signaling cells in the immune continuum*. Trends Immunol, 2004. **25**(9): p. 489-95.
42. Tomer, A., *Human marrow megakaryocyte differentiation: multiparameter correlative analysis identifies von Willebrand factor as a sensitive and distinctive marker for early (2N and 4N) megakaryocytes*. Blood, 2004. **104**(9): p. 2722-7.
43. Deutsch, V.R. and A. Tomer, *Megakaryocyte development and platelet production*. Br J Haematol, 2006. **134**(5): p. 453-66.
44. Patel, S.R., J.H. Hartwig, and J.E. Italiano, Jr., *The biogenesis of platelets from megakaryocyte proplatelets*. J Clin Invest, 2005. **115**(12): p. 3348-54.
45. Italiano, J.E., Jr. and R.A. Shivdasani, *Megakaryocytes and beyond: the birth of platelets*. J Thromb Haemost, 2003. **1**(6): p. 1174-82.
46. Italiano, J.E., Jr., et al., *Blood platelets are assembled principally at the ends of proplatelet processes produced by differentiated megakaryocytes*. J Cell Biol, 1999. **147**(6): p. 1299-312.
47. Richardson, J.L., et al., *Mechanisms of organelle transport and capture along proplatelets during platelet production*. Blood, 2005. **106**(13): p. 4066-75.
48. Reed, G.L., *Platelet secretory mechanisms*. Semin Thromb Hemost, 2004. **30**(4): p. 441-50.
49. Dobrovolsky, A.B. and E.V. Titaeva, *The fibrinolysis system: regulation of activity and physiologic functions of its main components*. Biochemistry (Mosc), 2002. **67**(1): p. 99-108.
50. Collen, D. and H.R. Lijnen, *Thrombolytic agents*. Thromb Haemost, 2005. **93**(4): p. 627-30.
51. Thom, T., et al., *Heart disease and stroke statistics--2006 update: a report from the American Heart Association Statistics Committee and Stroke Statistics Subcommittee*. Circulation, 2006. **113**(6): p. e85-151.
52. Wilkerson, W.R. and D.C. Sane, *Aging and thrombosis*. Semin Thromb Hemost, 2002. **28**(6): p. 555-68.

53. Cattabeni, F., F. Colciaghi, and M. Di Luca, *Platelets provide human tissue to unravel pathogenic mechanisms of Alzheimer disease*. Prog Neuropsychopharmacol Biol Psychiatry, 2004. **28**(5): p. 763-70.
54. Gnatenko, D.V., et al., *Transcript profiling of human platelets using microarray and serial analysis of gene expression*. Blood, 2003. **101**(6): p. 2285-93.
55. Denis, M.M., et al., *Escaping the nuclear confines: signal-dependent pre-mRNA splicing in anucleate platelets*. Cell, 2005. **122**(3): p. 379-91.
56. McRedmond, J.P., et al., *Integration of proteomics and genomics in platelets: a profile of platelet proteins and platelet-specific genes*. Mol Cell Proteomics, 2004. **3**(2): p. 133-44.
57. O'Neill, E.E., et al., *Towards complete analysis of the platelet proteome*. Proteomics, 2002. **2**(3): p. 288-305.
58. Marcus, K., et al., *Identification of platelet proteins separated by two-dimensional gel electrophoresis and analyzed by matrix assisted laser desorption/ionization-time of flight-mass spectrometry and detection of tyrosine-phosphorylated proteins*. Electrophoresis, 2000. **21**(13): p. 2622-36.
59. Garcia, A., et al., *Extensive analysis of the human platelet proteome by two-dimensional gel electrophoresis and mass spectrometry*. Proteomics, 2004. **4**(3): p. 656-68.
60. Martens, L., et al., *The human platelet proteome mapped by peptide-centric proteomics: a functional protein profile*. Proteomics, 2005. **5**(12): p. 3193-204.
61. Coppinger, J.A., et al., *Characterization of the proteins released from activated platelets leads to localization of novel platelet proteins in human atherosclerotic lesions*. Blood, 2004. **103**(6): p. 2096-104.
62. Moebius, J., et al., *The human platelet membrane proteome reveals several new potential membrane proteins*. Mol Cell Proteomics, 2005. **4**(11): p. 1754-61.
63. Maguire, P.B., et al., *Identification of the phosphotyrosine proteome from thrombin activated platelets*. Proteomics, 2002. **2**(6): p. 642-8.
64. Marcus, K., J. Moebius, and H.E. Meyer, *Differential analysis of phosphorylated proteins in resting and thrombin-stimulated human platelets*. Anal Bioanal Chem, 2003. **376**(7): p. 973-93.

2 Biomarkers and Biomarker Discovery

2.1 Overview

A Biomarker is defined as a characteristic, that is objectively measured and evaluated as an indicator of a normal biological process, a pathogenic process or pharmacologic response to a therapeutic intervention [1].

Focussed to clinical and diagnostic application, biomarkers are used to

- identify patients with a disease or abnormal condition
- classify the extent of a disease and monitor the progression
- give a prognosis
- predict and monitor the success of a therapeutic intervention

Besides physical properties like blood pressure, any type of molecule can be a biomarker. Transcriptional profiling and analysis of DNA methylation is frequently used for searching cancer biomarkers [2] and metabolomics can be used to find biomarkers in drug and toxicity studies [3]. Proteins are believed to be most ubiquitously affected in disease, response and recovery [4], so numerous studies are performed to discover biomarkers using proteomic approaches.

The perfect biomarker would be present or absent only if a certain condition occurs. Considering the complexity and diversity of biological systems, combined with detection limits of analysis methods, it is not likely that perfect biomarkers for diseases, which appear heterogeneous themselves, can be found. Thus, usually a threshold or range of abundance for a biomarker substance is defined in order to distinguish between the conditions of biological systems.

With the development of the “-omics” methods, it has become possible to separate and simultaneously quantify up to thousands of different species. These quantification data are processed using statistical methods to evaluate the species’ power to distinguish between different conditions or groups of individuals. Methods like cluster analysis or multivariate statistics are used to find patterns of species which allow differentiation between the biological conditions. Unfortunately, the fact of

simultaneous quantification introduces the problem of multiplicity and together with a disadvantageous ratio of the number of quantified species to the number of samples, it is hard to separate true positives and true negatives from false negatives and false positives [5]. A major reason for this uncertainty is the total variability experienced between unique measurements. The technical variability being one component of the total variability can be decreased to a minimum and kept at a constant level by standardisation of sampling and analysis procedures. The biological variability (variability between the individuals) is a result of heterogeneity in age, sex, lifestyle and disease state and can only be influenced (but also heavily biased) by patient selection.

2.2 Evaluation of potential biomarkers

Usually far more biomarker candidates emerge the initial discovery studies than are applicable to use in a diagnostic assay, even if it is planned to use a panel of markers rather than a single marker. A number of candidates can be eliminated prior to evaluation studies, because their change in abundance is known to happen also in other diseases or is related to conditions that apply only to one group like medical treatment or physical inactivity of critically ill patients. The key criterion of a biomarker candidate is its ability to distinguish between the tested groups, it is usually expressed in terms of selectivity and specificity.

Selectivity is the candidate's ability to identify the condition when it is present. Specificity is the candidate's ability to identify the absence of the condition when it is absent. The correlation of sensitivity and specificity is visualised in receiver-operator-characteristic (ROC) plots (sensitivity plotted against 1-selectivity).

When performing statistical tests, usually the significance level α (probability for false positive decision, $\alpha=1$ -sensitivity) is set to a certain value (mostly $\alpha=0.05$ or $\alpha=0.01$). The probability of a false negative decision β ($\beta=1$ -selectivity) depends on the value set for α (reduction of α means rise of β) and usually is not known.

The statistical power ($\text{power}=1-\beta$ =selectivity) depends on the chosen value for α , the variance of the data and the number of samples. Setting values for α , β and a desired expression difference and knowing the standard deviations of each dataset, it is possible to determine the number of samples needed.

The search for differentially expressed species can be interpreted as a series of independent test which all have the chance of α to be false positives. When performing n tests, the chance of finding no false positive drops from $1-\alpha$ for a single test to $(1-\alpha)^n$ (multiplicity problem). The solution is to reduce the value of α for the single tests (Bonferroni-correction [6] or Benjamini-Hochberg method [7]) to keep the desired level of α for the whole experiment.

2.3 Guidelines for biomarker discovery studies

A proteomic analysis represents only the first step in a biomarker discovery study, it would be part of the first phase (exploration phase) as suggested in the five phase scheme for the development of screening biomarkers for cancer proposed by the Early Detection Research Network (EDRN) [8]. This scheme provides guidelines which will largely apply to other biomarker discovery studies.

The selection of individuals, who form the groups which are used to search for biomarker candidates or to validate biomarkers, heavily depends on the aims of the study. So, in some designs, it can be better to compare the group of diseased individuals to a group of individuals with a similar disease instead of healthy individuals [9].

Phase I: Exploration phase

The aim of this phase is to establish a ranked list of biomarker candidates.

In this phase usually “-omics” methods are used to locate the biomarker candidates within expression profiles. The number of individuals in the compared groups usually is rather low (some ten individuals per group), the number of tested species high, which results in low confidence for these candidates.

In the exploration phase the material for biomarker searches is often directly derived from the affected tissues via biopsy to reduce the chance of interference and keep the concentration of possible biomarker substances as high as possible.

It is strongly suggested to perform an additional validation study in this phase using a new set of samples. This validation is best performed using another method like measuring protein expression levels for candidates found by RNA profiling or using immunological methods like ELISA, Western blot or protein arrays [10].

In this phase pooled samples are frequently used to keep the costs and sample numbers down, but pooling can lead to a loss in information by pushing species unique to few individuals below the detection limit [11].

Phase II: Clinical Assay Development

The aim of this phase is to establish an assay that is based on specimens that can be obtained non-invasively.

The biomarker candidates in this phase should be able to distinguish between diseased and healthy subjects. The candidate's separation power (true positive rate, false positive rate and ROC) is evaluated.

Phase III: Retrospective Screening

The power of the candidate to detect the disease before clinical diagnosis is evaluated in this phase.

The samples have been collected from diseased subjects before the clinical diagnosis of their disease.

Phase IV: Prospective Screening

In this phase the operating characteristics (detection rate, false referral rate) of the candidates when screening relevant population is determined.

An individual, that is tested positive in this phase, receives further diagnostic treatment.

Phase V: Control Studies

The aim of this phase is the estimation of the reduction in mortality rate in the screened population.

2.4 Influence of sample origin on biomarker discovery

The samples for biomarker discovery studies can be derived from different sources: Cells from cell culture and primary cells obtained from organisms via blood drawing or biopsy. To keep the variability introduced during the sample preparation procedure as low as possible, it is necessary to obtain a defined number of homogenous cells derived from a distinct morphological origin.

Sample preparation out of cell culture cells usually involves only the separation of cells from the medium they have been grown in. Obtaining cells out of tissues, the separation of the cells of interest from other cell types and the surrounding matrix is necessary. In case of blood, it is usually sufficient to use separation techniques based on the cells' physical properties (mass, density) to separate the major cell types. The isolation of cells out of solid tissues is a major challenge and often involves time and labour intense techniques like laser capture microdissection [12, 13] which allows the procurement of selected cells out of tissue sections.

Samples which are obtained from cell culture will be as homogenous as possible. These cells share the same genetic information, are grown under specific and controlled conditions and usually are available in almost unlimited amount. Often these cell lines are derived from cancer cells, which makes them a good model system for investigating cancer. These model systems can help to reduce the sample complexity, but unfortunately they will never show all properties of the real disease.

Studies based on cells isolated out of tissues obtained from organisms often suffer from the limited amount of material and contamination with different cell types. Samples derived from organisms usually show considerable variation between individuals caused by genetic heterogeneity, different lifestyle, simultaneous presence of other diseases and heterogeneous manifestation of the considered disease. Nevertheless, biomarker candidates that are discovered out of this material are more likely to pass the following validation phases because they were found by analysing samples from patients not model systems and have already shown to work in samples showing considerable variability.

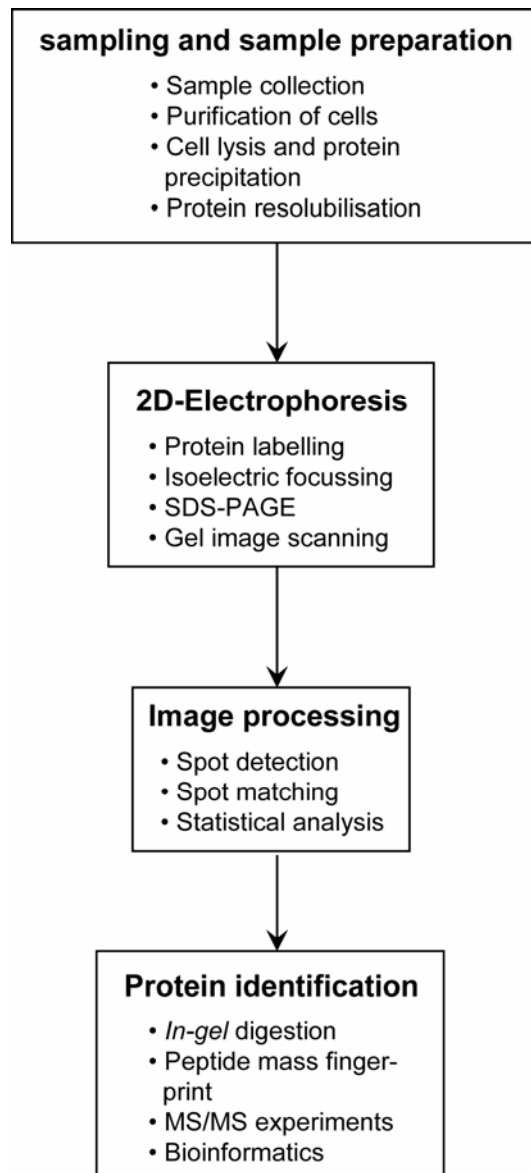
2.5 References

1. *Biomarkers and surrogate endpoints: preferred definitions and conceptual framework.* Clin Pharmacol Ther, 2001. **69**(3): p. 89-95.
2. Ramaswamy, S. and C.M. Perou, *DNA microarrays in breast cancer: the promise of personalised medicine.* Lancet, 2003. **361**(9369): p. 1576-7.
3. Fernie, A.R., et al., *Metabolite profiling: from diagnostics to systems biology.* Nat Rev Mol Cell Biol, 2004. **5**(9): p. 763-9.

4. Rifai, N., M.A. Gillette, and S.A. Carr, *Protein biomarker discovery and validation: the long and uncertain path to clinical utility*. Nat Biotechnol, 2006. **24**(8): p. 971-83.
5. LaBaer, J., *So, you want to look for biomarkers (introduction to the special biomarkers issue)*. J Proteome Res, 2005. **4**(4): p. 1053-9.
6. Sachs, L., *Angewandte Statistik*. 11th Edition ed. 2004: Springer.
7. Benjamini, Y. and Y. Hochberg, *Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing*. Journal of the Royal Statistical Society B, 1995. **57**(1): p. 289-300.
8. Pepe, M.S., et al., *Phases of biomarker development for early detection of cancer*. J Natl Cancer Inst, 2001. **93**(14): p. 1054-61.
9. Normolle, D., M. Ruffin, and D. Brenner, *Design of Early Validation Trials of Biomarkers*. Cancer Informatics, 2005. **1**(1): p. 25-31.
10. Cahill, D.J., *Protein and antibody arrays and their medical applications*. J Immunol Methods, 2001. **250**(1-2): p. 81-91.
11. Zolg, W., *The proteomic search for diagnostic biomarkers: lost in translation?* Mol Cell Proteomics, 2006. **5**(10): p. 1720-6.
12. Emmert-Buck, M.R., et al., *Laser capture microdissection*. Science, 1996. **274**(5289): p. 998-1001.
13. Niyaz, Y., et al., *Noncontact laser microdissection and pressure catapulting: sample preparation for genomic, transcriptomic, and proteomic analysis*. Methods Mol Med, 2005. **114**: p. 1-24.

3 Sources of Variation in the Biomarker Discovery Workflow

3.1 Workflow Description



Two-dimensional gel electrophoresis (isoelectric focussing (IEF) followed by SDS-polyacrylamide gel electrophoresis (SDS-PAGE)) was first described by O'Farrell [1]. The introduction of immobilised pH gradients for IEF [2] and their adaptation to 2D-electrophoresis [3] highly improved usability and reproducibility of this method. With

the introduction of differential gel electrophoresis (DIGE) [4] and the usage of an internal standard to minimise gel-to-gel variation [5], 2D-electrophoresis has become an accurate protein separation and quantification method.

Protein identification out of 2D-gels starts with excision of suitably stained protein spots [6], followed by *in-gel* digestion [7] with trypsin. The resulting peptide mixtures are analysed by matrix assisted laser desorption/ionisation mass spectrometry (MALDI-MS) [8-10] to derive the components' mass/charge (m/z) ratios. The derived m/z values depend on the protein's sequence (peptide mass fingerprint) and therefore are unique for each protein. Search engines compare submitted values to values that are generated by *in-silico* digestion of protein sequences and return a list of sequences ranked by a probability based score [11-13]. Peptide mass fingerprinting can be not sufficient for reliable protein identification, if two or more proteins are present in the same spot, or a high number of m/z values can not be explained by *in silico* digestion of the best ranked sequences. By using tandem mass spectrometry, it is possible to generate fragment ions from selected peptides either by post source decay (PSD) or collision-induced dissociation (CID), which show m/z values in dependence of the peptide's sequence. The obtained mass spectra can be used to perform *de novo* sequencing to generate sequence tags or the fragment ion's m/z values can be submitted to MS/MS search engines which compare these data with *in silico* generated MS/MS spectra and return a ranked list of peptides and corresponding proteins [11, 12].

3.2 Sources of Variation

During a workflow, that involves so many different steps, there is a variety of possible sources of variability.

Sampling and Sample Preparation

The variation, that is introduced in these steps, consist of the inevitable biological variation between individuals caused by factors like different genetic background, sex, age and lifestyle. Variations that are caused by neglecting circadian fluctuations in analyte concentration, drug treatment or presence of other diseases could be avoided by standardisation of the sampling procedure and proper patient selection.

During the process of sample preparation the main sources of variation are contamination, inadequate separation, material loss or degradation [14]. The standardisation of the sample preparation procedure should keep the variation introduced during this step as low as possible and at a constant level.

2D-Electrophoresis

The 2D-electrophoresis involves several steps where sample manipulation by an operator and the exposure of the sample to the environment is necessary. Zhou et al. [15] have investigated the amount of protein losses during this procedure. Artefacts in the spot pattern, mainly streaking, can be caused by insufficient protein solubility during isoelectric focussing, protein oxidation or salts [16]. Gels which have to be removed from the cassette, where they were cast in, to be stained and scanned, are easily damaged and distorted leading to differences in spot localisation. The selection of the staining method, also introduces variation [17], because dyes can affect over- or under-representation of certain protein spots due to the protein's physicochemical properties and each dye has its own sensitivity and dynamic range.

Image Analysis

Spot detection algorithms often fail to detect spots which are not completely separated or their volume drops below a certain intensity. Often the spot detection quality differs across a single gel with changing background levels and areas of bad resolution [18]. The more variability is created during spot detection, the more faulty matches are produced by the spot matching algorithms resulting in whole areas where all spots are mismatched. Wrong matching can be corrected and partially avoided by interference of operators, but this is time consuming and hard to reproduce. All errors in spot detection and spot matching influence the accuracy of the quantification results.

Protein Identification

Protein identification by peptide mass fingerprinting as well as sequence tags obtained from MS/MS experiments rely on the availability of a sequence database for the organism. Without the availability of sequence databases one has to perform *de novo* sequencing within MS/MS spectra and perform homology searches. The confidence of a hit relies on the quality of mass spectra (instrument resolution, limit of

detection and accuracy), enzyme specificity, the number of detected peptides and the quality of the sequence databases. Frequently, more than a single protein can be detected within a single spot [19] which renders any quantitation result obtained from this spot unusable. Also, peptides derived from a keratin contamination of the gel matrix or surface can reduce the confidence of the identification results or even completely mask the protein contained in the spot.

3.3 References

1. O'Farrell, P.H., *High resolution two-dimensional electrophoresis of proteins*. J Biol Chem, 1975. **250**(10): p. 4007-21.
2. Bjellqvist, B., et al., *Isoelectric focusing in immobilized pH gradients: principle, methodology and some applications*. J Biochem Biophys Methods, 1982. **6**(4): p. 317-39.
3. Gorg, A., W. Postel, and S. Gunther, *The current state of two-dimensional electrophoresis with immobilized pH gradients*. Electrophoresis, 1988. **9**(9): p. 531-46.
4. Unlu, M., M.E. Morgan, and J.S. Minden, *Difference gel electrophoresis: a single gel method for detecting changes in protein extracts*. Electrophoresis, 1997. **18**(11): p. 2071-7.
5. Alban, A., et al., *A novel experimental design for comparative two-dimensional gel analysis: two-dimensional difference gel electrophoresis incorporating a pooled internal standard*. Proteomics, 2003. **3**(1): p. 36-44.
6. Shevchenko, A., et al., *Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels*. Anal Chem, 1996. **68**(5): p. 850-8.
7. Jensen, O.N., et al., *Sample preparation methods for mass spectrometric peptide mapping directly from 2-DE gels*. Methods Mol Biol, 1999. **112**: p. 513-30.
8. Karas, M., et al., *Matrix-assisted ultraviolet laser desorption of non-volatile compounds*. International Journal of Mass Spectrometry and Ion Processes, 1987. **78**: p. 53.
9. Karas, M. and F. Hillenkamp, *Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons*. Anal Chem, 1988. **60**(20): p. 2299-301.
10. Tanaka, K., et al., *Protein and polymer analyses up to m/z 100 000 by laser ionization time-of-flight mass spectrometry*. Rapid Communications in Mass Spectrometry, 1988. **2**(8): p. 151-153.

11. Clauser, K.R., P. Baker, and A.L. Burlingame, *Role of accurate mass measurement (+/- 10 ppm) in protein identification strategies employing MS or MS/MS and database searching*. Anal Chem, 1999. **71**(14): p. 2871-82.
12. Perkins, D.N., et al., *Probability-based protein identification by searching sequence databases using mass spectrometry data*. Electrophoresis, 1999. **20**(18): p. 3551-67.
13. Zhang, W. and B.T. Chait, *ProFound: an expert system for protein identification using mass spectrometric peptide mapping information*. Anal Chem, 2000. **72**(11): p. 2482-9.
14. Castellanos-Serra, L. and D. Paz-Lago, *Inhibition of unwanted proteolysis during sample preparation: evaluation of its efficiency in challenge experiments*. Electrophoresis, 2002. **23**(11): p. 1745-53.
15. Zhou, S., et al., *A quantitative investigation into the losses of proteins at different stages of a two-dimensional gel electrophoresis procedure*. Proteomics, 2005. **5**(11): p. 2739-47.
16. Gorg, A., et al., *The current state of two-dimensional electrophoresis with immobilized pH gradients*. Electrophoresis, 2000. **21**(6): p. 1037-53.
17. Miller, I., J. Crawford, and E. Gianazza, *Protein stains for proteomic applications: which, when, why?* Proteomics, 2006. **6**(20): p. 5385-408.
18. Wheelock, A.M. and A.R. Buckpitt, *Software-induced variance in two-dimensional gel electrophoresis image analysis*. Electrophoresis, 2005. **26**(23): p. 4508-20.
19. Gygi, S.P., et al., *Evaluation of two-dimensional gel electrophoresis-based proteome analysis technology*. Proc Natl Acad Sci U S A, 2000. **97**(17): p. 9390-5.

Aims of this PhD Thesis

The aim of this PhD thesis was to establish a workflow for biomarker research in primary patient samples with focus on cells from the myeloid lineage.

The work is divided into three parts:

- Establishment of the biomarker discovery workflow (sample isolation and preparation => 2D gel electrophoresis => quantification of the intact proteins and statistical analysis => spot selection and protein identification by mass spectrometry) on the basis on a study with a monocytic cell line.
- Optimisation and validation of the preparation of primary patient material.
- Characterisation of the technical and biological variation in human blood samples.

Published Papers and Manuscripts

1 Quantitative validation of different protein precipitation methods in proteome analysis of blood platelets

Electrophoresis 2005; 26 (12): 2481 – 2489

Maria Zellner¹
Wolfgang Winkler¹
Hubert Hayden¹
Michael Diestinger¹
Maja Eliassen¹
Bernd Gesslbauer²
Ingrid Miller³
Martina Chang⁴
Andreas Kungl²
Erich Roth¹
Rudolf Oehler¹

¹Surgical Research Laboratories,
Medical University of Vienna,
Vienna, Austria

²Institute of Pharmaceutical
Sciences,
University of Graz,
Graz, Austria

³Department of Natural Sciences,
Institute of Medical Chemistry,
University of Veterinary Medicine,
Vienna, Austria

⁴Institute of Applied Microbiology,
University of Natural Resources
and Applied Life Science,
Vienna, Austria

Quantitative validation of different protein precipitation methods in proteome analysis of blood platelets

For the preparation of proteins for proteome analysis, precipitation is frequently used to concentrate proteins and to remove interfering compounds. Various methods for protein precipitation are applied, which rely on different chemical principles. This study compares the changes in the protein composition of human blood platelet extracts after precipitation with ethanol (EtOH) or trichloroacetic acid (TCA). Both methods yielded the same amount of proteins from the platelet preparations. However, the EtOH-precipitated samples had to be dialyzed because of the considerable salt content. To characterize single platelet proteins, samples were analyzed by two-dimensional fluorescence differential gel electrophoresis. More than 90% of all the spots were equally present in the EtOH- and TCA-precipitated samples. However, both precipitation methods showed a smaller correlation with nonprecipitated samples (EtOH 74.9%, TCA 79.2%). Several proteins were either reduced or relatively enriched in the precipitated samples. The proteins varied randomly in molecular weight and isoelectric point. This study shows that protein precipitation leads to specific changes in the protein composition of proteomics samples. This depends more on the specific structure of the protein than on the precipitating agent used in the experiment.

Keywords: Ethanol / Platelets / Protein precipitation / Proteomics / Trichloroacetic acid

DOI 10.1002/elps.200410262

1 Introduction

Proteome analysis of clinical samples aims at characterizing disease-specific changes in the protein expression profile of an affected tissue. Such changes are potential diagnostic markers or may help to identify drug targets. Blood platelets are small enucleated cellular particles that play a fundamental role in hemostasis, contributing to the formation of vascular plugs. Pathologically, they are involved in thrombosis and atherosclerosis. A detailed analysis of the proteome and signaling cascades in platelets from patients is expected to aid the development of new therapeutic agents that may help to treat thrombotic diseases [1, 2]. Platelets can be isolated from peripheral blood in a few preparative steps to a high cellular purity. To prevent activation of platelets, it is advisable to avoid strong centrifugation and to separate platelets from plasma protein by gel filtration. The gel-

filtered platelets are more quiescent, and are generally in better condition than platelets prepared by multiple steps of centrifugation [3]. However, it is impossible to concentrate the platelets using this technique. Therefore, the protein concentration is low and direct proteomic analysis of this platelet isolation is not possible. In addition, cells contain proteases and high levels of non-protein impurities. An appropriate sample preparation is essential for obtaining reliable results in a proteomic analysis. Such a preparation should work by inhibiting protease activities and quantitatively enriching proteins, while leaving behind substances, such as salts, lipids, and nucleic acids, which would interfere with any further proteomic analysis. Protein precipitation followed by dissolving the pellet in IEF compatible sample solution is generally employed to concentrate and selectively separate proteins in the sample from the interfering substances. In addition, the protein denaturation during precipitation leads to an inhibition of proteases [4].

The precipitation of proteins from biological fluids had been observed for hundreds of years (e.g., the precipitation of casein from milk by dilute acid). However, the molecular basis of protein solubility began to receive serious attention only in the middle of the last century. Fractionation of human plasma by precipitation carried forward

Correspondence: Professor Rudolf Oehler, Surgical Research Laboratories, General Hospital Vienna, Medical University of Vienna, Waehringer Guertel 18–20, A-1090 Vienna, Austria

E-mail: rudolf.oehler@meduniwien.ac.at

Fax: +43-1-40400-6782

Abbreviations: DIGE, differential gel electrophoresis; EtOH, ethanol; GFP, gel-filtered platelet; PRP, platelet-rich plasma; TM3, tropomyosin α 3

during World War II made possible the preparation of many protein components of the human plasma [5]. The solubility of proteins is determined by four variables: pH, ionic strength, temperature, and protein concentration. Numerous different techniques have been developed for protein precipitation by modifying one or more of these parameters (for review see [6]). Several methods have been adapted to the needs of proteome analysis [4, 7, 8] and today various protein extraction kits are commercially available which apply precipitation techniques. The present study evaluates the applicability of two methods of protein precipitation for clinical proteomics: ethanol (EtOH) precipitation and TCA precipitation. Both methods are commonly used in the preparation of protein extracts for proteomic analysis, although they rely on different chemical principles. EtOH causes precipitation of proteins mainly because it significantly lowers the dielectric constant of the aqueous solution (relative dielectric constant at 20°C for H₂O is 18 and for C₂H₅OH is 26). In general, ionic compounds are more soluble in solvents with high dielectric constants. Through its polar groups, EtOH interacts with the polar group of the protein in competition with water. In addition, the hydrophobic groups may disrupt the intramolecular hydrophobic interaction. Finally, a large volume of EtOH reduces the effective concentration of water, leaving only a small amount for hydration of the protein. Upon dehydration by EtOH, protein molecules attract each other to a sufficient degree by van der Waals forces and thus become insoluble in the EtOH-water mixtures [9]. TCA, in contrast, leads to a strong decrease in pH, resulting in denaturation and consequently precipitation of the protein. A recent study showed that the three chloro groups in the molecule also play an important role in protein precipitation, which is not clear [10]. TFA, which is a stronger acid than TCA and possesses three fluoro groups instead of chloro groups as in TCA, is not such a potent protein precipitation-inducing agent.

The precipitation efficiency of both EtOH and TCA depends also on the physicochemical characteristics of the protein. Therefore, they are also applied for the fractionation of protein mixtures. Especially EtOH is broadly used for protein fractionation even at industrial scale. Proteomics comprises the analysis of thousands of different protein species simultaneously. Applying EtOH or TCA precipitation in the analysis of such complex protein mixtures may result in depletion of particular protein species and in a relative enrichment of other protein species. The present study investigates qualitative and quantitative changes in the protein composition of platelet extracts after protein precipitation. It describes the protein yield, the linearity, and the protein selectivity of EtOH and TCA precipitation. By comparing these character-

istics with nonprecipitated samples, the study evaluates the suitability of protein precipitation in the proteome analysis of clinical samples of this type.

2 Materials and methods

2.1 Blood sampling and platelet preparation

Peripheral venous blood was drawn without stasis from 60 healthy volunteers aged between 20 and 90 years (30 female/30 male; average age 53 ± 21 years). The study was approved by the local ethics committee. Blood was drawn from the antecubital vein of the subjects into vacutainer tubes containing 0.129 mol/L trisodium citrate (Vacurette system; Greiner, Kremsmuenster, Austria). The first 3 mL of blood was discarded as usual for platelet studies. For exclusion of erythrocytes and leukocytes, the citrated whole blood was centrifuged at 50 × *g* for 20 min at room temperature. The resulting supernatant was the platelet-rich plasma (PRP). A plasma-free platelet suspension was prepared by passing PRP through a size-exclusion chromatography (SEC) column. One milliliter PRP was applied onto 11 mL packed Sepharose 2B (Sigma, Steinheim, Germany) column (BioRad, Hercules, CA, USA; 15 mm diameter) equilibrated in calcium-free Dubecco's PBS (GIBCO, Paisley, Scotland, UK). Platelet fractions (1.5 mL) of each individual were collected after an elution volume of 2.5 mL, and platelet concentration was counted on a MicroDiff 18 Blood Analyzer (Coulter Electronics, Miami, FL, USA).

2.2 TCA precipitation

Fifteen-hundred microliters of platelet suspension was mixed with 500 μL of ice-cold 6.1 N TCA solution (Sigma) containing 80 mM DTT (Roche Diagnostics, Mannheim, Germany). The mixture was incubated for 1 h at 4°C to allow the protein precipitation to complete. Then the extract was centrifuged at 10 000 × *g* for 10 min at 4°C. The supernatant was discarded, and the pellet was washed four times with 1500 μL of ice-cold acetone (p.a. grade; Merck, Darmstadt, Germany) each, containing 20 mM DTT; the pellet was regained in each step by centrifugation at 10 000 × *g*. Thereafter, the centrifuged pellet was dried by air evacuation. For 2-DE, the pellet was resolubilized in denaturing 2-D sample buffer containing 7 M urea, 2 M thiourea, 4% CHAPS, 30 mM Tris-HCl (pH 8.5) by shaking overnight at 4°C. Seventy microliters of the sample buffer was used *per* 100 × 10⁶ platelets. Alternatively, the samples were stored in the last wash aliquot of acetone at -70°C until further use.

2.3 EtOH precipitation

Nine-hundred microliters of platelet suspension was mixed with 8000 μL EtOH (99.9% Uvasol p.a.; Merck). Proteins were allowed to precipitate in this EtOH solution for 3 h at room temperature followed by storage at -20°C overnight. To collect the proteins, the samples were centrifuged at $10\,000 \times g$ for 20 min at 4°C and the supernatant was removed. The pellet was washed once with pure EtOH before it was dried (as above). For 2-DE, the pellet was resolubilized in 100 μL of 2-D sample buffer as above *per* 100×10^6 platelets. Alternatively, the samples were stored in the 90% EtOH solution at -20°C until further use.

2.4 Dialysis

To reduce the salt contamination of 2-D samples from EtOH-precipitated platelet proteins, they were dialyzed against a $40 \times$ sample volume of identical 2-D sample buffer for 3 h at room temperature with PlusOne Mini Dialysis Kit (molecular mass cut-off 1 kDa) (Amersham Biosciences, Uppsala, Sweden).

2.5 Preparation of proteins from nonprecipitated platelets

Fifteen-hundred microliters of gel-filtered platelet (GFP) suspension was centrifuged at $1500 \times g$ for 10 min at room temperature. The supernatant was discarded and the platelet pellet was solubilized in 100 μL of 2-D sample buffer as above *per* 100×10^6 platelets.

2.6 Determination of protein concentration

The protein concentration in resolubilized samples was determined in triplicate using a CBB protein assay kit with BSA as the standard protein (Pierce Biotechnology, Rockford, IL, USA). With appropriate predilution, the 2-D sample buffer components do not interfere with the protein assay. Therefore, the samples were diluted 1:20 with PBS and 5% of the 2-D sample buffer was added in the BSA standards.

2.7 Analysis of platelet proteins by 1-DE and 2-DE

IEF was performed loading 120 μg (for gels to be silver-stained) or 150 μg (for gels with CyDye-labeled proteins) of platelet proteins by in-gel rehydration in a volume of 450 μL , denaturing the 2-D buffer (7 M urea, 2 M thiourea, 4% CHAPS, 70 mM DTT, 0.5% Servalyt™ pH 3–10;

Serva, Heidelberg, Germany) onto 24 cm IPG DryStrips, pH 3–10 linear and pH 4–7 linear (Amersham Biosciences) and focused for 50 and 30 kVh, respectively, using an Amersham IPGphor unit. Before loading onto SDS-polyacrylamide gels, IPG strips were incubated for 15 min in equilibration buffer (50 mM Tris-HCl, pH 8.8, 6 M urea, 30% glycerol, 2% SDS) containing 1% DTT and then for another 15 min in equilibration buffer containing 2.5% iodoacetamide. The SDS-polyacrylamide gels (26 \times 20 cm \times 1 mm, T = 11%, C = 2.6%) were cast according to Laemmli [11]. The second dimension was performed using an Ettan DALT six System (Amersham Biosciences) according to the manufacturer's instructions. 1-D SDS-gel electrophoresis was performed with 13 \times 16 cm gels.

2.8 Protein staining and image analysis of silver-stained gels

Silver nitrate staining for analytical gels was performed according to Heukeshoven and Dernick [12], while gels for MS analysis were stained with mass-compatible silver stain according to Shevchenko [13]. The silver-stained gel images were digitized using a Molecular Imager® FX (BioRad). Computer-aided 2-D image analysis was carried out using the MELANIE 3 software (GeneBio, Geneva, Switzerland).

2.9 Protein labeling and image analysis of DIGE gels

Fifty micrograms of resolubilized platelet protein preparations was labeled with 333 pmol of CyDye DIGE Fluor minimal dyes (Amersham Biosciences). Pre-electrophoretic labeling was performed according to the manufacturer's instructions. The gels with separated labeled proteins were scanned using the Typhoon™ 9410 imager (Amersham Biosciences), and the protein patterns were displayed with the IQTools software. All sample gel images were processed by the DeCyder DIA (differential in-gel analysis) software (Amersham Biosciences) module to codetect and differentially quantify the protein spots.

2.10 In-gel protein digestion

The spots of interest were excised from the gels, chopped into pieces, and transferred into 0.5 mL tubes (Axygen, Union City, CA). The gel pieces were washed twice with 200 μL of 50 mM NH_4HCO_3 buffer (pH 8.5) (Sigma) for 10 min and afterwards with 200 μL of 50 mM NH_4HCO_3 in 50% ACN (HPLC-grade; Merck) for 10 min. Subsequently, the gel pieces were dehydrated by adding 50 μL ACN and

allowed to reswell in 180 μL of 10 mM DTT in 50 mM NH_4HCO_3 buffer in order to perform reduction (56°C, 30 min). After cooling to room temperature, the solution was replaced by 150 μL of 50 mM iodoacetamide in 50 mM NH_4HCO_3 buffer and the gel pieces were incubated in the dark for 20 min at room temperature. The gel pieces were washed three times with 200 μL of 50 mM NH_4HCO_3 buffer and three times with 50 mM NH_4HCO_3 in 50% ACN for 10 min at room temperature, dehydrated with 50 μL ACN, and dried in the Eppendorf Concentrator 5301 (Eppendorf, Hamburg, Germany) for 5 min. The enzymatic digestion of the proteins was carried out on ice by a stepwise addition of 0.5–2 μL of 12.5 ng/ μL trypsin (sequencing-grade unmodified; Roche, Basel, Switzerland) in 50 mM NH_4HCO_3 buffer until they were totally rehydrated. Finally, enough 50 mM NH_4HCO_3 buffer was added to keep the gel pieces covered during digestion at 37°C overnight. After digestion, the supernatant was removed and the peptides were extracted once with 20 μL 50 mM NH_4HCO_3 buffer and twice with 20 μL 5% formic acid (Sigma) by sonification for 5 min at room temperature in an ultrasonic water bath (Sonorex RK 255 H; Bandelin, Berlin, Germany).

2.11 NanoHPLC-MS/MS protein sequencing

All nanoHPLC separations were performed on the Ultimate system from LC Packings (Amsterdam, The Netherlands). The in-gel digests were loaded onto a precolumn (PepMap C18 material, 300 μm ID \times 5 mm length; LC Packings) by the FAMOS μ -autosampler and the Switchos loading pump operated at 20 $\mu\text{L}/\text{min}$ using water with 0.1% TFA (Pierce, Rockford, IL, USA) as mobile phase. The sample was eluted from the precolumn in a back flush mode. The dimensions of the separation column were 0.075 mm ID \times 150 mm length, 3 μm particle size. The flow rate of the nanoHPLC system was set at 200 nL/min and the UV detector was operated at 214 nm using the nano UV-Z view flow cell (volume 3 nL). The mobile phases were A = 95% water (HPLC-grade, Supra-Gradient, Biosolve B.V., The Netherlands), 5% ACN (HPLC-grade, Supra-Gradient, Biosolve B.V.), 0.1% formic acid (Fluka, Buchs, Switzerland); and B = 30% water, 70% ACN, 0.1% formic acid. The HPLC gradient for separation was 0–50% B in 30 min and 50–100% B in 2 min. The nanoHPLC system was coupled to an IT mass spectrometer (LCQ Deca XPplus, Thermo Finnigan) via a nanoESI source using Pico Tip emitters (New Objective, Cambridge, MA, USA). The following ESI parameters were used: spray voltage, 1.8 kV; capillary temperature, 185°C; capillary voltage, 45 V; tube lens offset voltage, 25 V; and the electron multiplier at –1050 V. The collision energy was set automatically depending on the mass of the parent ion.

Gain control was set to 5×10^7 . The data were collected in the centroid mode using Dynamic Exclusion. One MS experiment (full-MS) was followed by three MS/MS experiments of the three most intensive ions (intensity at least 1×10^6). The analysis of MS/MS spectra with respect to the peptide identity was routinely performed by applying both the MASCOT (Matrix Science) and the SEQUEST (Thermo Finnigan) search engines. A peptide was reliably identified only if the individual peptide scores were >43 (MASCOT) and >3.2 (SEQUEST).

3 Results

3.1 Sample quality of precipitated platelets and compatibility with 2-DE

To evaluate and compare the applicability of TCA precipitation and EtOH precipitation for clinical proteomics, we used these methods for the preparation of proteins from blood platelets for 2-DE. GFP showed a mean concentration of $117 \pm 42 \times 10^6$ Plt/mL. This suspension was subjected to a precipitation either with TCA or with EtOH, according to the procedures described in Section 2. The 2-D gel of TCA-precipitated proteins showed a clear protein pattern with about 1400 individual spots, which were found with good reproducibility (Fig. 1A). However, EtOH-precipitated proteins were only well separated in the lower *pI* range (Fig. 1B). In the *pI* range above 6.0, the proteins were only seen as horizontal streaks. This indicates a nonsufficient separation of EtOH-precipitated proteins in the first-dimensional IEF. Figures 1D–F show the electrical current and the voltage profile of the IEF separation. The high current peak in the first part of the IEF of EtOH-precipitated proteins indicates that this sample contained high salt concentrations (Fig. 1E, arrow). These salts had to be removed by dialysis in order to get a satisfactory separation by 2-DE (Fig. 1C and F). This dialysis step was included later on in all following EtOH-precipitation experiments.

3.2 Protein yield

The protein precipitation methods were evaluated in the GFP preparations of 60 different volunteers. The amount of proteins, which was extracted from 100×10^6 platelets after TCA precipitation, was equal to that extractable after EtOH precipitation from the same number of platelets (Fig. 2A). The protein amount increased linearly with the number of platelets used in the experiments (Fig. 2B). As an alternative approach for increasing the platelet concentration of GFP, we centrifuged the cell suspension and extracted the proteins from the pellet in the 2-D sample

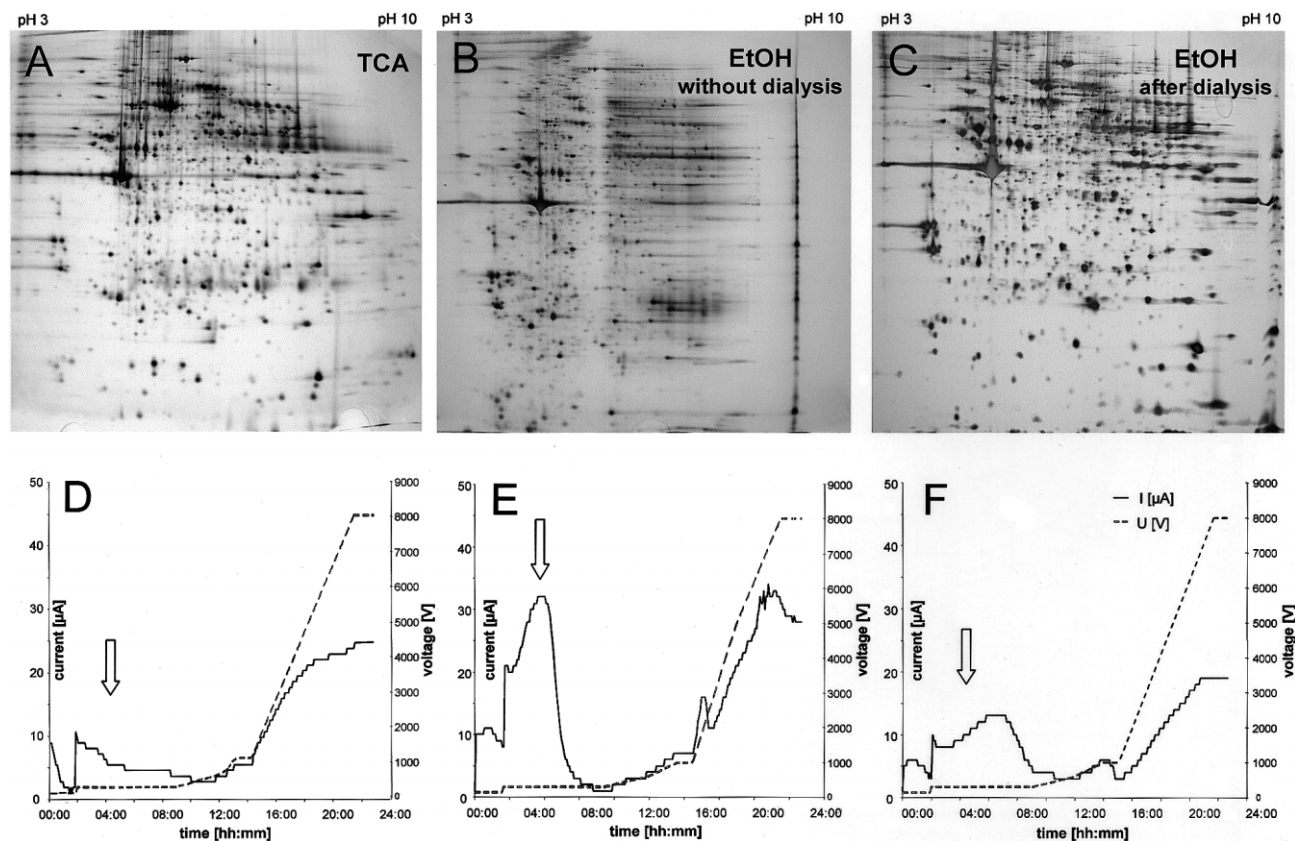


Figure 1. 2-D electrophoretic analysis of gel-filtered human platelet suspension following precipitation with (A) TCA and (B) EtOH resolubilized in 2-D sample buffer. In (C) EtOH-precipitated and -solubilized platelet proteins were dialyzed against fresh 2-D sample buffer to remove salt contamination. Twelve-hundred microliters of prepared platelet proteins was loaded by passive rehydration onto 24 cm pH 3–10 IPG strips. Current and voltage profile were recorded of (D) TCA-, (E) EtOH-, and (F) EtOH-precipitated and -dialyzed platelet samples during the 1-D. Gels were stained with analytical silver nitrate stain.

buffer. The supernatant contained almost no protein (2 μg proteins *per* 100×10^6 platelets). The amount of proteins extracted from these nonprecipitated platelets was similar to that found after precipitation (124 ± 10 μg proteins *per* 100×10^6 platelets). These data show that both precipitation methods have a similar protein yield and indicate that precipitation followed by resolubilization is not accompanied by a major protein loss.

3.3 Specificity of protein precipitation

To characterize potential protein specific differences in precipitation efficiency, we performed a comparative electrophoretic analysis. Nonprecipitated, TCA-precipitated, and EtOH-precipitated platelet proteins were separated in 1-D SDS-PAGE. The results are shown in Fig. 3A. The protein band pattern is similar in all three samples. Only the nonprecipitated platelets show some slight bands in the upper molecular weight range, which

are barely detectable in the precipitated samples. TCA- and EtOH-precipitated proteins showed a nearly identical pattern of protein bands.

For definitive comparison of different precipitation methods on platelet protein pattern, we switched to the differential gel electrophoresis (DIGE) technology. TCA-precipitated proteins, EtOH-precipitated proteins, and nonprecipitated proteins were labeled with three different fluorescent dyes, combined and separated by 2-DE. The overlay of the three signals is shown in Fig. 3B. Most protein spots are equally present in all three samples (black spots). However, some protein spots which are present in the nonprecipitated samples are diminished in the TCA-precipitated and EtOH-precipitated samples (magenta spots). These proteins are equally distributed all over the 2-D gel. Almost no TCA-specific protein spots (blue) or EtOH-specific protein spots (yellow) are visible. However, some green protein spots are present in the 2-D gel, indicating proteins which are enriched by both precipitation

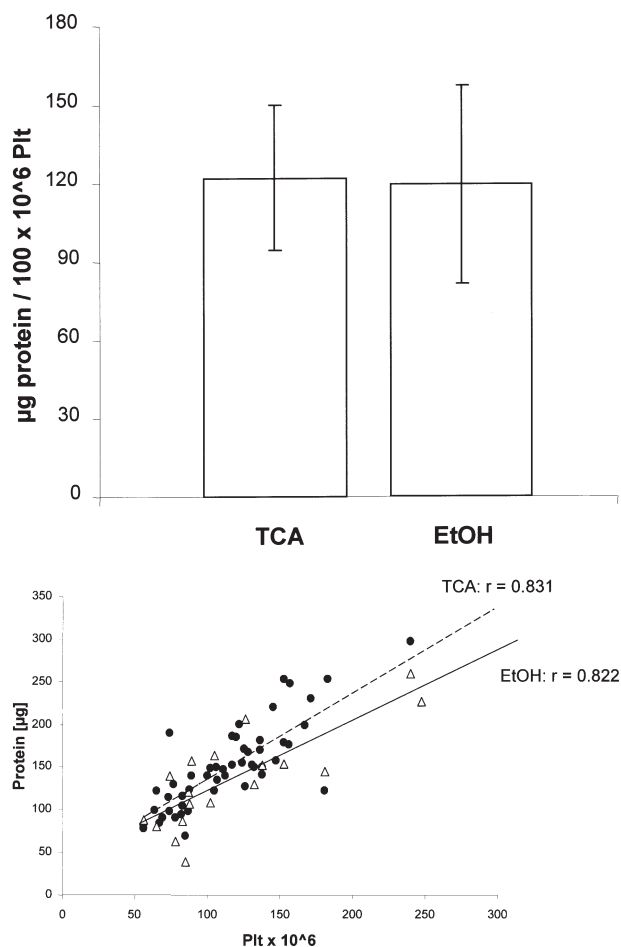


Figure 2. Platelet protein yield for each precipitation method. (A) GFP preparations of different volunteers ($n = 18$) were precipitated with TCA as well as with EtOH and resolubilized in 2-D sample buffer ($70 \mu\text{L}/100 \times 10^6$ platelets for TCA and $100 \mu\text{L}/100 \times 10^6$ platelets for EtOH precipitation). EtOH-precipitated samples were dialyzed. Protein concentration was determined by Bradford. Values are expressed as mean value \pm SD. (B) Platelet protein yield for each precipitation method of differently concentrated GFP suspensions. GFP preparations of different volunteers were precipitated with TCA ($n = 51$ volunteers; closed circle) or with EtOH ($n = 18$ volunteers; open triangle). Linear regression was calculated for protein yield of differently concentrated GFP suspensions precipitated by TCA or EtOH.

methods. Proteins which showed a modified abundance in precipitated samples in at least two independent experiments are indicated in Fig. 3B and their identity is shown in Table 1.

To quantify the protein-specific properties of the three extraction methods a scatter plot analysis was performed. Therefore, we plotted the fluorescence intensity signal of every single spot of the EtOH- or TCA-pre-

cipitated sample against the nonprecipitated sample (Figs. 3C and D) and of the EtOH-precipitated sample against the TCA-precipitated sample (Fig. 3E). About 75% of all protein spots are present at the same levels in the EtOH-precipitated samples and nonprecipitated samples, whereas 11.6% are over-represented in the EtOH samples and 13.5% in nonprecipitated samples (Fig. 3C). Similarly, 79% are present at the same level in the TCA-precipitated samples and nonprecipitated samples (Fig. 3D). However, when comparing the EtOH-precipitated samples with the TCA-precipitated samples, 91% of all protein spots turned out to be equally present in both samples (Fig. 3E). This scatter graph shows a substantially narrower Gaussian distribution of the protein spots than the other two graphs. This indicates a much stronger similarity between the two precipitation methods than between one single precipitation method and the nonprecipitated samples. Almost all of the 8.7% over-represented in TCA samples were overlapping spots within the two spikes left and right of the β -actin spots (chains of light blue spots indicated by arrows in Fig. 3B). Since β -actin was found in the whole spike, it could not be determined whether this overrepresentation is derived from other proteins or is a β -actin effect.

4 Discussion

The present study evaluates the usability of two different protein precipitation methods, the EtOH precipitation and TCA precipitation, for the proteomic analysis of human platelets. It characterizes the precipitation efficiency on an overall as well as on a protein-specific level.

Due to the distinct chemistry of EtOH and TCA, the protein samples are treated in a different way for precipitation. For EtOH precipitation, platelets were diluted in a 9-fold volume of EtOH and then incubated at room temperature. Studies with the human plasma showed that most proteins precipitate already at an EtOH concentration of 40% v/v [5]. However, there were still some soluble proteins found under such conditions. The high surplus of EtOH used in our study facilitates the precipitation of all proteins. EtOH is readily miscible with water, but yields a significant heat of solution and has the tendency to denature proteins, especially at temperature above 0°C [6]. In protein fractionation experiments, which are aiming at preparing nondenatured proteins, normally a “cold EtOH precipitation” protocol is used, in which the solvent temperature is always kept below 0°C . In proteomics, however, denaturation is unproblematic as far as it is not accompanied by protein modification. It is, therefore, not necessary to decrease the temperature during EtOH precipitation. In addition, EtOH has a lower dielectric con-

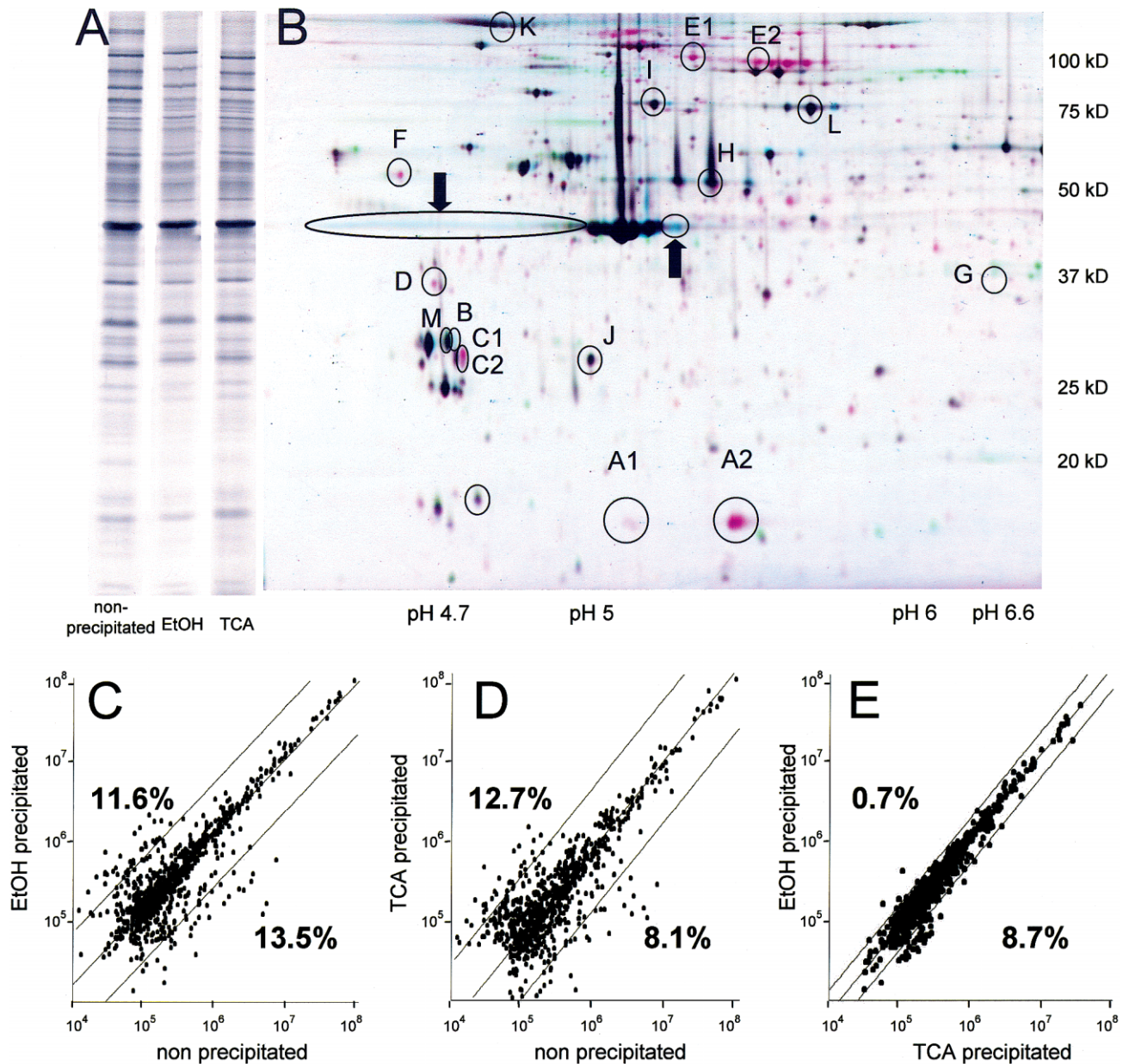


Figure 3. 1-D and 2-D DIGE-analysis of nonprecipitated, TCA- and EtOH-precipitated platelet samples. (A) Silver-stained 1-D gel with 2 μ g protein from each platelet preparation. (B) 2-D fluorescence overlay image of 50 μ g protein from each simultaneously made platelet preparation are focused in one IPG strip (pH 4–7). TCA-precipitated platelet proteins are shown in blue, EtOH-precipitated platelet proteins in yellow, and nonprecipitated platelet proteins in magenta. Proteins which are equally present in each platelet protein preparation are displayed in black. The figure shows one of two independent experiments. (C–E) Scatter plot analysis of the 2-D gel shown in (B). Each dot represents one protein spot. Its position indicates the fluorescence intensity signal in the respective preparation group. Analysis includes 829 spots, which were detected in all preparation groups in two independent experiments. Regression line and confidential intervals indicating the twofold SD are shown. R indicates the linear correlation coefficient. (C) Nonprecipitated *versus* EtOH-precipitated samples; $R = 0.988$, (D) nonprecipitated *versus* TCA-precipitated samples; $R = 0.975$, (E) EtOH-precipitated *versus* TCA-precipitated samples; $R = 0.979$.

Table 1. Identifications of representative precipitation-dependent and -independent proteins

Spot ID	Protein name	Accession number
Protein spots decreased in TCA- and EtOH-precipitated platelet samples		
A1	Transgelin 2 isoform	Q9BUH5
A2	Transgelin 2 isoform	Q9BUH5
C1	TM3 ^{a)} chain isoform	P06753
C2	Myosin heavy chain β -subunit	Q14905
D	TM3, fibroblast	NCBI 88928
E1	Vinculin isoform	P18206
E2	Vinculin isoform	P18206
F	Nucleosome assembly protein 1-like 1	P55209
Protein spots increased in TCA and EtOH-precipitated platelet samples		
B	TM3 chain isoform	P06753
G	PDZ and LIM domain	O00151
Protein spots equally present in all platelet samples		
H	Fibrinogen γ chain A	P02672
I	Heat shock cognate 71 kDa protein	P11142
J	Chloride channel	O00299
K	Integrin α IIb isoform	P08514
L	Albumin	P02768
M	TM3 chain isoform	P06753

a) Tropomyosin α 3

stant at room temperature than below 0°C, which further facilitates the precipitation [9]. The platelets used in this study were suspended in a buffered isotonic salt solution in order to avoid hypotonic lysis. Due to the low solubility of salts in alcohols, EtOH precipitation is not suitable to remove salts from the platelet sample, which resulted in an insufficient protein separation in the IEF. This effect was also observed when the precipitation was performed at -20°C (data not shown). To reduce the salt concentration, the resolubilized precipitate had to be dialyzed which is a time and cost intensive extension of the EtOH-precipitation protocol. However, the quantitative validation of the EtOH-precipitated and dialyzed samples indicate that dialysis has almost no effects on the protein composition of the sample.

For TCA precipitation, the precipitant was added to the platelet suspension to a final concentration of 25%. The TCA-induced protein precipitation curves are observed to be U-shaped [10]. All proteins are soluble at very low TCA concentrations and the first precipitation of serum albumin can be observed when the concentration increases

above 5%. Between 15 and 40% TCA even the highly TCA-soluble protein cardiotoxin III precipitates. When the acid concentrations are raised above 50%, all proteins are found to redissolve back into the solution. This is thought to be mediated by acid-induced structural transitions in the proteins. Such irreversible protein modification should be avoided in the preparation of samples for proteome analysis. Thus, the TCA concentration used in the present study is suited to precipitate almost all proteins with a minimum of protein modification. In order to avoid any acid-induced protein hydrolysis during storage, the precipitate is centrifuged and washed in acetone to remove TCA. The organic solvent, acetone, is on one hand, a protein precipitant with a similar dielectric constant as EtOH, and on the other hand, a good solvent of TCA.

The present study shows that, in spite of the different chemistries and the different precipitation protocols, EtOH and TCA precipitation lead to similar results in the proteomic analysis. Both the methods seem to induce precipitation of almost all proteins of the platelet suspension. The incubation of the precipitated proteins in the strongly denaturing 2-D buffer, which contained chaotropic agents and detergents, resolubilized almost all proteins. The protein content of this solution was independent of the used precipitant. The amount of protein found after precipitation and resolubilization was the same as in a 2-D buffer extract from the same number of platelets, which were concentrated by centrifugation instead of precipitation. However, both precipitation methods lead to changes in the protein composition of the sample comprising the reduction and increase of specific proteins. The amount of proteins depleted by the precipitation was compensated by a similar amount of proteins, which were enriched by this treatment. All mammalian cells and platelets, in particular, are very sensitive against mechanical stress during cell preparation. Platelets become highly activated by strong centrifugal forces, which lead to artificial and badly reproducible modifications of the proteome [3]. Concentration of the gel-filtered platelets by centrifugation at high speed should, therefore, be avoided. Precipitation instead stabilizes the proteome of unstressed cells and, in addition, denatures any protease activities. The precipitation and resolubilization induced changes in the protein composition were found to be highly reproducible. Therefore, this preparation procedure was preferred. Unfortunately, from our results it cannot be estimated whether the depletion of proteins is due to an incomplete precipitation or resolubilization, and whether the enrichment is due to higher precipitation efficiency or to a higher solubility of these proteins in the 2-D buffer.

The EtOH precipitation and TCA precipitation lead to depletion of the same proteins (summarized in Table 1). These include actin associated proteins such as transgelin 2 (SM22 α), tropomyosin, myosin heavy chain, and vinculin. These proteins span a broad *pI* and molecular weight range and are equally distributed throughout the gel. This indicates that there is no preferential precipitation or resolubilization of large or small, acidic or basic proteins. The selectivity of precipitation and resolubilization for certain protein species might derive from a different hydrophobicity of these proteins. However, calculating the grand average of hydrophobicity (GRAVY) score [14] revealed no difference between the group of depleted and the group of unchanged proteins (data not shown). Thus, the selectivity must rely on other structural features of the proteins. TM3 is of special interest. Vertebrate nonmuscle cells, such as human platelets, express multiple isoforms of tropomyosin, which are generated from four different genes and a combination of alternative promoter activities and alternative splicing [15]. In the present study, various TM3 isoforms behave differently in precipitation: while one isoform (Fig. 3B, spot C1/C2) is decreased after precipitation, another isoform (spot B) is increased, and a third isoform (M) remains unchanged. Over 20 different TM3 isoforms have been identified so far. All TM3 isoforms conform to a general pattern of structure. Their amino acid sequences consist of an integral number, six or seven in vertebrates, of quasiequivalent regions of about 40 residues that are considered to represent the actin-binding regions of the molecule. In addition to the variable regions, a large part of the polypeptide chains of the TM3 isoforms is invariant. Thus, TM3 isoforms differ only slightly in their structure. Both *in vitro* and *in vivo* evidence suggest that multiple isoforms of TM3 in non-muscle cells may be required for regulating actin filament stability, intracellular granule movement, cell shape determination, and cytokinesis [16]. These various functions are carried out by binding to different proteins, such as caldesmon, tropomodulin, and other unidentified proteins. Since the precipitation experiments were performed in the current study with intact platelets, these binding partners might affect the precipitation and subsequent resolubilization properties of the TM3 isoforms leading thereby to depletion or enrichment of the protein.

Taken together, the current study shows that protein precipitation leads to specific changes in the protein composition of a proteomics sample. These changes are more influenced by the specific features of the protein than by the precipitant used in the experiment.

We thank Martin Ploder for help with the statistic analysis, Susanne Oehler for helpful discussion, and Hemma Bauer for technical advice.

Received October 16, 2004

5 References

- [1] Garcia, A., Prabhakar, S., Hughan, S., Anderson, T. W., Brock, C. J., Pearce, A. C., Dwek, R. A., Watson, S. P., Hebestreit, H. F., Zitzmann, N., *Blood* 2004, 103, 2088–2095.
- [2] Maguire, P. B., Wynne, K. J., Harney, D. F., O'Donoghue, N. M., Stephens, G., Fitzgerald, D. J., *Proteomics* 2002, 2, 642–648.
- [3] McNicol, A., in: Watson, S. P., Authi, K. S. (Eds.), *Platelets: A Practical Approach*, Oxford University Press, New York 1996, pp. 1–26.
- [4] Rabilloud, T., Chevallet, M., in: Rabilloud, T. (Ed.), *Proteome Research: Two-dimensional Gel Electrophoresis and Identification Methods*, Springer, Berlin 2000, pp. 9–27.
- [5] Cohn, E. J., Strong, L. E., Hughes, W. L., Mulford, D. J., Melin, M., Taylor, H. L., Ashworth, J. N., *J. Amer. Chem. Soc.* 1946, 68, 459–475.
- [6] England, S., Seifter, S., *Methods Enzymol.* 1990, 182, 285–300.
- [7] Jiang, L., He, L., Fountoulakis, M., *J. Chromatogr. A* 2004, 1023, 317–320.
- [8] Nandakumar, M. P., Shen, J., Raman, B., Marten, M. R., *J. Proteome Res.* 2003, 2, 89–93.
- [9] van Oss, C. J., *J. Prot. Chem.* 1989, 8, 661–668.
- [10] Sivaraman, T., Kumar, T. K., Jayaraman, G., Yu, C., *J. Prot. Chem.* 1997, 16, 291–297.
- [11] Laemmli, U. K., *Nature* 1970, 227, 680–685.
- [12] Heukeshoven, J., Dernick, R., *Electrophoresis* 1988, 9, 28–32.
- [13] Shevchenko, A., Wilm, M., Vorm, O., Mann, M., *Anal. Chem.* 1996, 68, 850–858.
- [14] Kyte, J., Doolittle, R. F., *J. Mol. Biol.* 1982, 157, 105–132.
- [15] Perry, S. V., *J. Muscle Res. Cell. Motil.* 2001, 22, 5–49.
- [16] Lin, J. J., Warren, K. S., Wamboldt, D. D., Wang, T., Lin, J. L., *Int. Rev. Cytol.* 1997, 170, 1–38.

2 Adaptive cellular mechanisms in response to glutamine-starvation

Frontiers in Bioscience 2006; 11: 3199 – 3211

Adaptive cellular mechanisms in response to Glutamine-starvation

Maja Munk Eliassen¹, Wolfgang Winkler^{1,2}, Veronika Jordan¹, Marion Pokar¹, Martina Marchetti², Erich Roth¹, Günter Allmaier² and Rudolf Oehler¹

¹ Department of Surgery - Research Laboratories, Medical University of Vienna and ²Institute of Chemical Technologies and Analysis, Vienna University of Technology, Vienna, Austria

TABLES OF CONTENT

1. Abstract
2. Introduction
3. Material and Methods
 - 3.1. Chemicals
 - 3.2. Cell culture and treatment
 - 3.3. Analysis of intracellular glutamine
 - 3.4. Western blot analysis
 - 3.5. Analysis of intracellular ATP
 - 3.6. Analysis of intracellular glutathione
 - 3.7. Cell lysis and determination of protein concentration
 - 3.8. Analysis of cell proteins by one and two dimensional gel electrophoresis
 - 3.9. Protein identification
4. Results
 - 4.1. Effect of long-term Gln-starvation on cellular metabolism
 - 4.2. Effect of long-term Gln-starvation on cell protective mechanisms
 - 4.3. Effect of long-term Gln-starvation on overall protein synthesis
 - 4.4. Effect of long-term Gln-starvation on the proteome
5. Discussion
6. Acknowledgment
7. References

1. ABSTRACT

Glutamine (Gln) utilising cells suffer from Gln-starvation during critical illness when plasma Gln levels are decreased. This study investigates whether such cells activate adaptive mechanisms. Monocytic U937 cells were cultured at 0.6 and 0.2 mM Gln for up to four days. Within the first day a decrease of ATP (78% of control), intracellular free Gln (13%), Hsp70 (74%) and proliferation rate (79%) was observed. A prolonged culture at 0.6 mM Gln for additional three days led to a recovery of ATP (97%), Hsp70 (91%) and proliferation (92%). The intracellular free Gln increased only to 41%. At 0.2 mM Gln, however, all levels remained decreased. The activation of the metabolic sensor AMP activated kinase (AMPK) increased immediately in Gln-starving cells but regained normal values only in cells cultured at 0.6 mM. A proteomic analysis identified 23 proteins, which were affected by Gln starvation including metabolic enzymes, proteins involved in synthesis and degradation of RNA and proteins, and stress proteins. These data show that Gln-utilising cells activate adaptive mechanisms in response to Gln-starvation, which enable them to overcome a Gln shortage. At very low Gln concentrations, these adaptive mechanisms are not sufficient to counteract the lack of the amino acid.

2. INTRODUCTION

With a concentration of 0.5-0.6 mM, glutamine (Gln) is the most abundant free amino acid in human plasma. Due to this high concentration, Gln is a ready source of energy and is used besides glucose as a supplementary energy substrate for the intestine and in particular for monocytes/macrophages. During critical illness, the Gln consumption-rate exceeds the supply, and plasma as well as skeletal muscle pools of free Gln are severely reduced. The plasma Gln levels can drop below 0.4 mM and monocytes suffer from Gln-starvation. Many organisms and tissues react to a reduced amino acid supply with activation of adaptive cellular mechanisms in order to overcome the shortage. The identified levels of regulation include transcriptional control, as well as mRNA stability and translational control. Numerous *in vitro* studies revealed that Gln-starvation induces many changes in phenotype and function of Gln-utilising cells. Gln depletion of primary monocytes and monocytic U937 cells for only 3 hours resulted in a reduction in proteasome-mediated proteolysis (1). This was accompanied by an accumulation of ubiquitin-protein conjugates and a reduction of intracellular ATP. Furthermore, U937 cells were more susceptible to the apoptotic triggers Fas-ligand and TNF- α 4 h after withdrawal of Gln (2). Studies in leukaemia-

Adaptation to glutamine starvation

derived CEM and HL-60 cells revealed that Gln-deprivation induces ligand-independent CD95 receptor signalling and apoptosis (3). In a recent publication, we could show that Gln-starvation of monocytic U937 cells for 4 h induces specific alterations in their proteome with the stress protein Hsp70 showing the strongest reduction (4). This decrease in Hsp70 expression was due to an enhanced decay of Hsp70 mRNA. All these studies clearly show that Gln-starvation induces a specific response immediately after withdrawal of the amino acid in primary monocytes as well as in monocytic U937 cells. However, it cannot be concluded whether these cells activate adaptive mechanisms in response to Gln, whether the described changes belong to such a mechanism and whether these cells are able to compensate for the lack of the amino acid. Some hints to an adaptive capacity of monocytic U937 cells come from earlier studies in which cells were exposed to Gln-starvation for up to five days (5). Although this treatment led to a reduced proliferation and an increased cell volume, the cells showed no impairment of their viability. Similarly, primary monocytes can be cultured in the absence of Gln for up to five days without any effect on their survival (6). Thus, these cells must somehow overcome the lack of Gln. The present study aims to investigate whether Gln-utilising cells are able to adapt to Gln-starvation and to characterise such a response. Therefore, we analysed monocytic U937 cells exposed to different degrees of Gln-starvation for up to four days. This long-term Gln-starvation approach allowed the analysis of the immediate cell response as well as of the succeeding adaptation to the reduced Gln supply. We could clearly show that U937 cells are able to compensate for mild Gln-starvation and could identify proteins involved in this process by a proteomic analysis.

3. MATERIAL AND METHODS

3.1. Chemicals

Radiochemicals were from Amersham Bioscience (Uppsala, Sweden), o-phthalaldehyde, potassium borate solution (pH = 10.4), and β -mercaptoethanol were purchased from Pierce (Rockford, IL, USA). All other chemicals used were obtained from Sigma (St. Louis, MO, USA) if not specified. All chemical were p.a. grade.

3.2. Cell culture and treatment

Human premonocytic cells U937 (CRL-1593.2, ATCC, Rockville, MD, USA) were cultured in RPMI-1640 medium (Gibco, Paisley, UK) supplemented with 10 % FCS and 2 mM Gln at 37°C in a humidified atmosphere of 95 % air and 5 % CO₂. For Gln-starvation the same medium but with reduced Gln-supplementation was used. Without any supplemented Gln the medium still contained 0.05 mM Gln deriving from FCS.

3.3. Analysis of intracellular glutamine

Cells were washed in PBS, counted, 5×10^6 cells were suspended in 100 μ l 3 % 5-sulfosalicylic acid containing 100 μ M β 2-thienylalanine and the suspension was shock-frozen and defrosted. After centrifugation, the amino acids in the supernatant were analysed by HPLC: 1 volume of OPA-reagent (13 mg o-phthalaldehyde, 200 μ l

methanol, 800 μ l 1 M potassium borate pH = 10.4, 25 μ l β -mercaptoethanol and 9 ml water) was added and 60 seconds later the sample was loaded onto a 150 x 3 mm reversed phase Hypersil-C₁₈ HPLC column (Maisch, Ammerbuch, Germany) on a Beckmann Gold HPLC-system (Beckmann, Palo Alto, CA, USA) fitted with a JASCO FP-920 fluorometer (JASCO, Tokyo, Japan). Separation was performed with a programmed gradient elution using 13 mM sodium acetate pH = 6.8, in 0.28 % tetrahydrofuran as buffer A and 50 % acetonitrile in 1.8 % tetrahydrofuran as buffer B (gradient: 0.1 min. 0 %B, 0.3 min. increase to 17.5 %B, 2.6 min.s 17.5 %B, 0.4 min. increase to 19.0 %B, 4.6 min. 19.0 %B, 0.5 min. increase to 25.0 %B, 4 min. 25.0 %B, 1.0 min. increase to 32.0 %B, 0.5 min. 32.0 %B, 7 min. increase to 43 %B, 3.6 min. 43 %B, 1.4 min. increase to 63.0 %B, 1 min. increase to 100 %B with a flow rate of 0.485 ml / min.). Fluorescence output ($\lambda_{\text{ex}} = 330$ nm; $\lambda_{\text{em}} = 408$ nm) was quantified by comparison to the internal standard (β 2-thienylalanine) as well as to data obtained for a standard mixture of amino acids.

3.4. Western blot analysis

Whole cell lysates were obtained by hypotonic treatment (10 mM Tris-HCl pH = 7.8, 1 mM EDTA, 10 mM KCl, 0.3 % Triton X-100, 1 mM phenylmethylsulfonyl fluoride, 1 μ g/ml leupeptin, and 1 μ g/ml aprotinin) combined with short sonification at 0°C. After centrifugation (10 min, 15,000 g, 4°C) the supernatant was stored in aliquots at -70°C. The cellular protein content was measured using a Bradford assay (Pierce, Rockford, IL, USA). For Western blot analysis, 10 μ g samples were separated by SDS-PAGE transferred to a nitrocellulose membrane by electroblotting. Hsp70 expression was revealed with an Hsp70 specific antibody (SPA-810, Stressgen, Victoria, Canada). The levels of AMPK and phospho AMPK were determined by specific antibodies (Upstate, Charlottesville, VA, USA).

3.5. Analysis of intracellular ATP

The intracellular ATP level was determined using the bioluminescence luciferin/luciferase assay system with modifications as previously described (7). Briefly, U937 cells were washed twice in PBS, counted and 1 ml 0.5 M HClO₄ was added to 10^6 cells. After centrifugation (14,000 g for 10 min) 50 μ l of supernatant was mixed with 12.5 μ l 1 M Na₂CO₃ to elevate the pH. Then 937.5 μ l TAE (100mMTris-acetate pH = 7.8, 2 mM EDTA) were added. Two hundred microliters of this solution was filled in a transparent vial and after automatic injection of 50 μ l ATP monitoring reagent (Bio-Orbit, Turku, Finland), light emission was measured for 20 s in a chemiluminometer (Auto-Lumat LB953, Berthold, Wildbad, Germany).

3.6. Analysis of intracellular glutathione

Cells were harvested and washed twice with ice cold PBS. Two x 10^6 cells were mixed vigorously with 1 ml of an aqueous 6.5 % sulfosalicylic acid (Merck, Darmstadt, Germany). After 15 min incubation on ice, the supernatant was snap frozen in liquid nitrogen and stored at -70°C until further analysis by HPLC. Hundred microliters of the neutralised sample or glutathione (GSH) standard

Adaptation to glutamine starvation

were mixed with 100 μ l MBB (0.57 % in acetonitril and sodium *N*-ethylmorpholine, pH 10.1), and allowed to react in the dark for 5 min before the reaction was stopped by the addition of 10 μ l 50 % sulfosalicylic acid (8). The HPLC separation of low molecular mass thiol-bimane adducts was achieved on a Supelcosil LC-18 octadecyl-silyl silica column (150 mm X 4.6 mm, 3 μ m particle size; Supelco, Bellefonte, PA, USA), followed by fluorimetric detection (EP920 fluorescence detector; Jasco, Osaka, Japan) at an excitation wavelength of 394 nm and an emission wavelength of 480 nm. Elution solvent A was composed of aqueous 9 % acetonitril, 0.25 % acetic acid and 0.25 % perchloric acid. The pH was adjusted to 3.7 with 40 % sodium hydroxide. Elution solvent B was based on a 75 % aqueous solution of acetonitril. The gradient elution consisted of 100 % A for 7 min., followed by 100 % B to elute matrix interferences, and returning to solvent A for re-equilibration for 12 min. at a flow rate of 1.0 ml/min. The resultant profiles were quantified on a basis of peak areas and compared with external standards of GSH (usually 0.625-10 μ mol/l).

3.7. Cell lysis and determination of protein concentration

Proteins were extracted and quantified as previously described (9). Briefly, 1500 μ l of cell suspension were mixed with 500 μ l of ice-cold 6.1 N TCA solution containing 80 mM DTT (Roche, Mannheim, Germany). The mixture was incubated for one hour at 4°C to allow protein precipitation to complete. Then the extract was centrifuged at 10,000 g for 10 min at 4°C. The supernatant was discarded and the pellet washed four times with 1,500 μ l of ice-cold acetone each, containing 20 mM DTT; the pellet was regained in each step by centrifugation at 10,000 g. Thereafter the centrifuged pellet was dried by air evacuation. The pellet was resolubilized in denaturing 2D sample buffer containing 7M urea, 2M thiourea, 4 % CHAPS, 30 mM Tris-HCl (pH=8.5) by shaking overnight at 4°C. The protein concentration in resolubilized samples was determined in triplicates using a Coomassie protein assay kit with BSA as the standard protein (Pierce, Rockford, IL, USA). With appropriate pre-dilution, 2D sample buffer components do not interfere with the protein assay. Therefore, samples were diluted 1:20 with PBS and 5% of 2D sample buffer were added in the BSA standards.

3.8. Analysis of cell proteins by one and two dimensional gel electrophoresis

For two-dimensional gel electrophoresis (2DE) proteins were separated in the first dimension by isoelectric focusing (IEF) and in the second dimension by SDS-polyacrylamide gel electrophoresis. IEF was performed loading 120 μ g of proteins by in-gel rehydration in a volume of 450 μ l denaturing 2D buffer (7 M urea, 2 M thiourea, 4% CHAPS, 70 mM DTT, 0.5 % servalyt™ pH 3-10; Serva, Heidelberg, Germany) onto 24 cm IPG DryStrip pH 3-7 linear (Amersham Bioscience, Uppsala, Sweden) and focused for 50 kVh using an Amersham IPGphor unit. Before loading onto SDS polyacrylamide gels, IPG strips were incubated for 15 min in equilibration buffer (50 mM Tris-HCl pH 8.8, 6 M urea, 30 % glycerol, 2 % SDS) containing 1 % DTT and then for another 15 min in

equilibration buffer containing 2.5% iodoacetamide. SDS polyacrylamide gels (260 x 200 x 1 mm, T = 11 %, C = 2.6 %) were cast according to Laemmli (10). The protein spots were visualised by a ruthenium II tris (bathophenanthroline disulfonate) staining (11). Comparative spot pattern analysis was accomplished with Delta-2D software (Decodon, Greifswald, Germany). One-dimensional SDS-gel electrophoresis was performed with 130 x 160 x 1.5 mm gels.

3.9. Protein identification

The spots of interest were excised manually using a stainless steel scalpel and subjected to in-gel digestion procedure (12) using trypsin (from bovine pancreas, modified; sequencing grade, Roche, Mannheim, Germany). The solutions of tryptic peptides, extracted out of the gel pieces, were desalted and purified utilising ZipTip® technology (C₁₈ reversed phase, standard bed Millipore, Bedford, MA USA) (13). The sample preparation for MALDI mass spectrometry was carried out on a stainless steel target, applying thin layer preparation technique (14) using α -Cyano-4-hydroxy-cinnamic acid (Prod-nr. C2020, Sigma Aldrich, St. Louis, MO, USA) as matrix dissolved in acetone (6 mg/ml). Positive ion mass spectra were recorded on a prototype vacuum MALDI TOF/curved field reflector TOF instrument (Shimadzu Biotech-Kratos Analytical, Manchester, UK) equipped with a nitrogen-Laser ($\lambda=337$ nm) by accumulating up to 500 single unselected laser shots. The instrument was operated in the reflectron mode, applying 20 kV acceleration voltage and delayed extraction (optimal setting for m/z 2000). External calibration was performed using an aqueous solution of standard peptides (Bradykinin fragment 1-7, human Angiotensin II, ACTH fragment 18-39). The lists of m/z-values derived from the mass spectra of in-gel digestions were submitted to the publicly accessible peptide mass fingerprint (PMF) search engines MS-Fit (15) and MASCOT (16) to search the SWISSPROT protein sequence database (Vers. 48.1 – 48.3) (17) and the non-redundant database from NCBI (Release 2005/06/01) (18) applying restrictions for species (*homo sapiens*), molecular weight (value estimated from 2D gel \pm 33%), isoelectric point (value estimated from 2D gel \pm 2) and mass tolerance (\pm 0.35 Da). A PMF search result was considered only to be a correct identification if the protein was rated as a significant match in the context of the reliability scoring system of both search algorithms. To confirm each identification result, post source decay (PSD) or low energy collision induced dissociation (CID) MS/MS experiments were performed on selected peptides. PSD experiments were performed on the high vacuum MALDI TOF/curved field reflector TOF instrument using the same sample preparation procedure as described previously. Low energy CID experiments were performed on a HCT^{plus} 3D-ion trap mass spectrometer (Bruker Daltonics, Bremen, Germany) equipped with a second generation atmospheric pressure (AP)-MALDI pulsed dynamic focusing (PDF)-source (Agilent Technologies, Palo Alto, CA, USA) and a nitrogen laser ($\lambda=337$ nm). The eluate from ZipTip® purification was prepared onto a gold coated target using the dried droplet technique (19) The eluate from ZipTip® purification was prepared onto a gold coated target using the dried

Adaptation to glutamine starvation

droplet technique (6 mg/ml). Positive ion AP-MALDI mass spectra were obtained using the following parameters: 350°C dry gas temperature and 100 ms accumulation time. Precursor ions for low energy CID-experiments were selected with an isolation width of 4 Da. External calibration was performed using the same standard peptide mixture as described above. The minimum requirement for successful confirmations of the PMF result was either at least one tryptic peptide yielding enough fragment ions to facilitate the manual assignment of a consecutive amino acid sequence, or two tryptic peptides, which did not allow the readout of a sequence tag, but reached significant MOWSE scores when submitting the list of m/z values to MASCOT MS/MS ion search (16). These searches have been performed using restrictions for species (*Homo sapiens*), database (SwissProt, NCBItr), precursor and product ion tolerance (PSD-data: ± 1 Da, low energy CID-data: ± 0.5 Da).

4. RESULTS

4.1. Effect of long-term Gln-starvation on cellular metabolism

To evaluate the effect of long term Gln-starvation on cellular metabolism cells were cultured in a medium supplemented with different Gln concentrations (2.0 mM, 0.6 mM, 0.2 mM, and 0.0 mM Gln) for up to 4 days. Since U937 cells show optimal growth at 2 mM Gln, this concentration was used as control. The intracellular level of free Gln (Gln_i) was measured before start of the experiment and at every single day. As shown in Figure 1A, at 2.0 mM Gln in the medium the intracellular Gln concentration was slightly below 6000 nmol Gln per μg protein. This level remained stable during the four days. In contrast, at reduced extracellular Gln the intracellular level dropped down immediately below 1000 nmol Gln per μg protein within 24 h. Only cells cultured in presence of 0.6 mM Gln were able to increase their Gln_i after 48 h to about 35% of control. However, there was no further increase observed in the following days in these cells. The Gln_i of cells cultured at 0.2 mM or 0.0 mM Gln remained below 5% of control for the whole experiment. These data clearly showed that extracellular Gln influenced strongly the intracellular concentration of free Gln. Only at 0.6 mM Gln, cells seemed to activate successful compensative mechanisms to increase their Gln_i. The next experiment investigated the effect of Gln-starvation on cell proliferation. Cells were cultured in medium with 2.0 mM, 0.6 mM, 0.2 mM, and 0.0 mM Gln for up to 96h. Cells were collected every 24 h, counted and seeded freshly with a density of 5×10^5 cells per ml. This helped to avoid any influence of cell density. The daily increase in cell number (daily replication rate) was used as a measure for cell proliferation with 1 indicating no proliferation. The results are shown in Figure 1B. At 2.0 mM Gln, the daily cell replication rate remained stable at a level slightly above 2.2. This indicated that cell number increased by this factor within 24h. At reduced Gln concentration, daily replication rate declined to about 1.6 within 24h. Cells cultured at 0.6 mM Gln returned to a normal daily replication rate within the following 24h. At lower Gln levels, the daily replication rate remained reduced for the rest of the experiment. It has to be mentioned that the daily replication rate was never below

1.4, even at 0.0 mM Gln. This indicated that also in the absence of Gln a cell proliferation occurred. Next, we investigated the effect of Gln-starvation on the intracellular concentration of ATP. Cells were cultured in medium with 2.0 mM, 0.6 mM, 0.2 mM, and 0.0 mM Gln for up to 96h and ATP was measured every day. As shown in Figure 1C, the availability of Gln had a strong influence on ATP level. At 2.0 mM Gln the ATP level remained unchanged during over 96h. At 0.6 mM Gln, ATP levels declined to 80% of normal within the first day and remained reduced until 72h. Then the ATP level increased to normal in the last day. At 0.2 mM Gln, ATP declined in the first day to 58% of normal and remained always below 68% throughout the experiment. At 0.0 mM Gln, ATP levels declined in the first day down to 44% of normal. In the following days, it decreased further with the lowest level at 32% of normal after 3 days. Taken together, these experiments show that the effect of Gln-starvation on cellular metabolism depends on the level of still available Gln. Cells seem to activate mechanisms for compensation of the missing amino acid. However, only cells exposed to mild Gln-starvation (0.6 mM Gln) are able to overcome the lack of Gln to a certain extent.

4.2. Effect of long-term Gln-starvation on cell protective mechanisms

Decreased intracellular ATP levels result in an increase of AMP, which is known to activate the AMP-activated kinase (AMPK) in a variety of cells by phosphorylation. Activated AMPK finally down-regulates energy consuming processes and stimulates ATP production to protect cells from starvation. To investigate whether Gln-starvation is able to activate AMPK we exposed cells to 2.0 mM, 0.6 mM, and 0.2 mM Gln for up to 96 h. Cells were harvested after 24 h and after 96 h and AMPK-phosphorylation was investigated by Western blotting. As Shown in Figure 2, Gln-starvation resulted in an increased phosphorylation of AMPK within 24 h. The level of phosphorylation depended on the degree of Gln-starvation. After 96 h the AMPK phosphorylation returned to normal in cells exposed to 0.6 mM Gln. In contrast, AMPK phosphorylation remained high in cells exposed to 0.2 mM Gln. In a recently published study, we could show that short-term Gln-starvation results in a specific reduction of stress protein Hsp70 expression by reduced mRNA stability (4). To evaluate the effect of long-term Gln-starvation on Hsp70 expression we cultured cells at 2.0 mM, 0.6 mM, and 0.2 mM Gln for 24 h and 96 h. Then Hsp70 expression was determined by Western blotting. As shown in Figure 3, Hsp70 expression declined at reduced Gln supply within 24 h to about 75% of normal. However, after 96 h Hsp70 expression was found to increase to normal levels in cells cultured at 0.6 mM Gln while it remained decreased at 0.2 mM.

4.3. Effect of long-term Gln-starvation on overall protein synthesis

Gln is, as an amino acid, a substrate for protein synthesis. Therefore, we investigated the effect of long-term Gln-starvation on protein synthesis rate and cellular protein content. Cells were cultured in medium with 2.0 mM, 0.6 mM, 0.2 mM, and 0.0 mM Gln for 24 h and

Adaptation to glutamine starvation

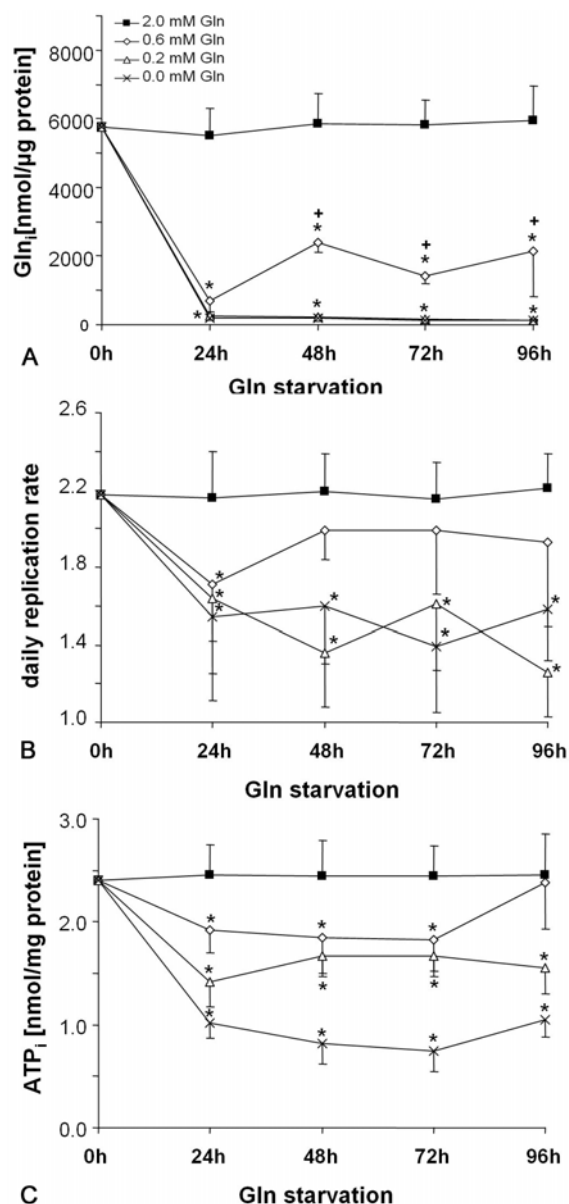


Figure 1. Influence of extracellular glutamine supplementation on cellular metabolism. U937 cells were cultured in medium supplemented with various concentrations of Gln for 24-96h. Cells were freshly seeded each day and density never exceeded 0.5×10^6 cells/ml. Apoptosis or necrosis were excluded by trypan blue stain (viability was always $\geq 95.6\%$). At each time point, cells were harvested and analysed. (A) Intracellular free Gln (Gln_i). Gln_i was measured by HPLC and set into relation to total cellular protein. Data are mean \pm SD of $n = 5$. (B) Daily replication rate. Cells were counted every day and the duplication rates within the last 24 h were calculated. Data are mean \pm SD of $n = 8$. (C) Intracellular ATP levels (ATP_i). ATP_i concentrations were determined by a luminescence assay and set into relation to total cellular protein. Data are mean \pm SD of $n = 5$. Statistical analysis is an one-way ANOVA with contrast against control (2mM Gln; significance indicated with an asterisk; Dunnett, 95% CI did not include 0) or against 0.6mM Gln (significance indicated with a cross).

96 h. To measure the cellular protein content cells were lysed, free amino acids were removed by TCA precipitation and protein amount in the precipitate was measured. As shown in Figure 4A, Gln-starvation had no effect on protein content of cells. To investigate the influence of Gln-starvation on the overall protein synthesis rate, cells were cultured under the same conditions as above and 2 h before end of the experiment [³⁵S]-Met was added for metabolic protein labelling. Then cells were lysed and proteins were separated in a SDS-gel electrophoresis. The autoradiogram of the gels indicated the [³⁵S]-Met incorporation into the proteins. We took the intensities of the autoradiogram between 20 kDa and 150 kDa as a measure of overall protein synthesis. As shown in Figure 4B, there was no effect of Gln-starvation on overall protein synthesis rate.

4.4. Effect of long-term Gln-starvation on the proteome

To evaluate whether Gln-starvation influences the expression of individual protein species without affecting the overall protein synthesis rate we used a proteomics approach. Cells were cultured in medium with 2.0 mM, 0.6 mM, and 0.2 mM Gln for 24 h and 96 h. Then proteins were extracted and separated by two-dimensional gel electrophoresis. Correspondingly, the resulting gels were termed: 2.0/24, 0.6/24, 0.2/24, 2.0/96, 0.6/96, and 0.2/96. Protein spots in the gel were visualised by ruthenium II tris (bathophenanthroline disulfonate) staining followed by digital fluorometric scanning. For quantification of spot intensities, the fluorescence signal intensities within the spot boundaries were used to calculate the fluorescence intensity volume (FIV) of each spot. An average of 595 spots (506-672) was detected in a typical gel (Figure 5). About ninety percent of these spots (532 spots) were found repeatedly in every 2.0 mM gel. Only these spots were used for further analysis. To minimise gel-to-gel variations, the FIV of every single spot in a gel was normalised by dividing it by the sum of all FIV values of all spots in the gel (FIV_{all}). The resulting percentage fluorescence intensity volume (%FIV) was used as a measure for protein spot abundance. The scatter blot in Figure 6 compares this abundance of all 535 protein spots in the 2.0/24 gel with their abundance in the 2.0/96 gel. The comparison nicely showed that all spots had nearly the identical abundance in both gels. This indicated that at 2.0 mM Gln the protein expression profile does not change during time. Therefore, the %FIV_{2.0/24} was used as normal control to reveal the effect of Gln-starvation on protein spot abundance. The %FIV values of gels from Gln-starved cells (0.6/24, 0.2/24, 0.6/96 or 0.2/96) were divided by the %FIV values of the corresponding spot in the control gel (%FIV_{2.0/24}). The log₂ value of these ratios visualises a stimulatory or inhibitory effect of Gln-starvation in comparison to control: a log₂ value of +1 indicates an increase by factor 2 and a log₂ value of -1 a decrease by factor 2, whereas a log₂ value of zero indicates no change (Figure 7). The log₂(ratios) are plotted against the spot abundance. Most protein spots group in the y-axis around zero. Their deviation from zero corresponds to the expected normal distribution. This clearly shows that they are unaffected by a Gln-starvation. In addition, the log₂(ratio) of these spots is independent of the spot abundance (x-axis). This excludes any artificial influence of the 2D gel

Adaptation to glutamine starvation

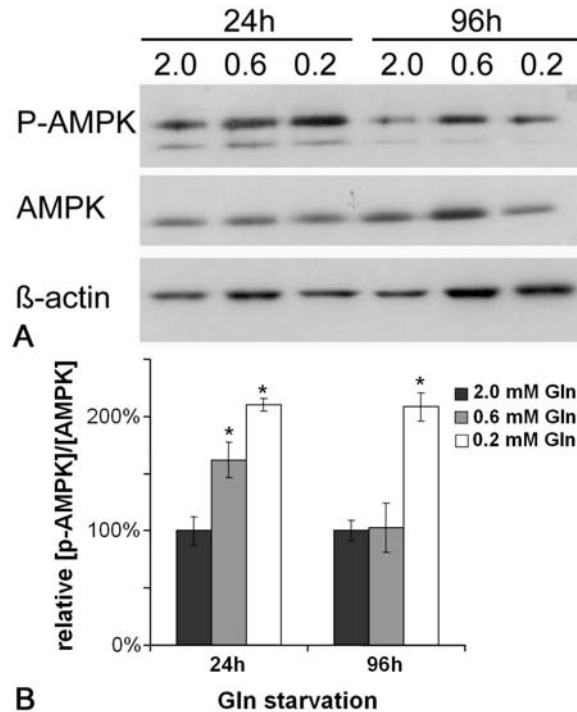


Figure 2. Influence of Gln-starvation on AMP-activated Protein Kinase (AMPK). U937 cells were cultured in medium supplemented with various concentrations of Gln for 24 and 96h. At both time point, cells were harvested and expression of AMPK as well as the level of phosphorylated AMPK (thr172, P-AMPK) was determined by Western blot. (A) Western blot analysis. The upper panel shows the expression of phosphorylated AMPK, the middle panel shows the level of whole AMPK, and the lower panel shows β -actin levels for loading control. (B) Statistical evaluation. The Western blot signal of five independent experiments were quantified by densitometry and statistically evaluated. AMPK activation is expressed as mean ratio of P-AMPK in relation to total expression of AMPK. The mean ratio of control cells (2.0 mM Gln) was set to 100%. The asterisks indicate $p < 0.05$ compared to control cells as calculated by Students t-test.

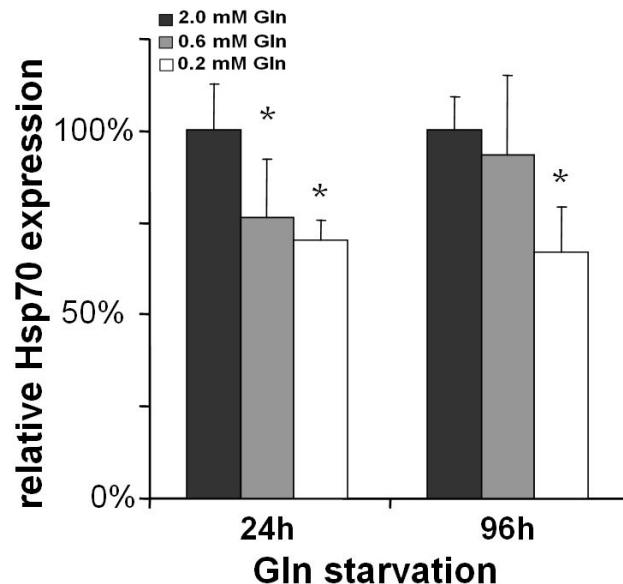


Figure 3. Influence of Gln-starvation on Hsp70 expression. U937 cells were cultured in medium supplemented with various concentrations of Gln for 24 and 96h. At both time points, were harvested and expression of Hsp70 was determined by Western blot. The mean Hsp70 in control cells (2.0 mM Gln) was set to 100%. The asterisks indicate $p < 0.05$ compared to control cells as calculated by Students t-test ($n = 3$).

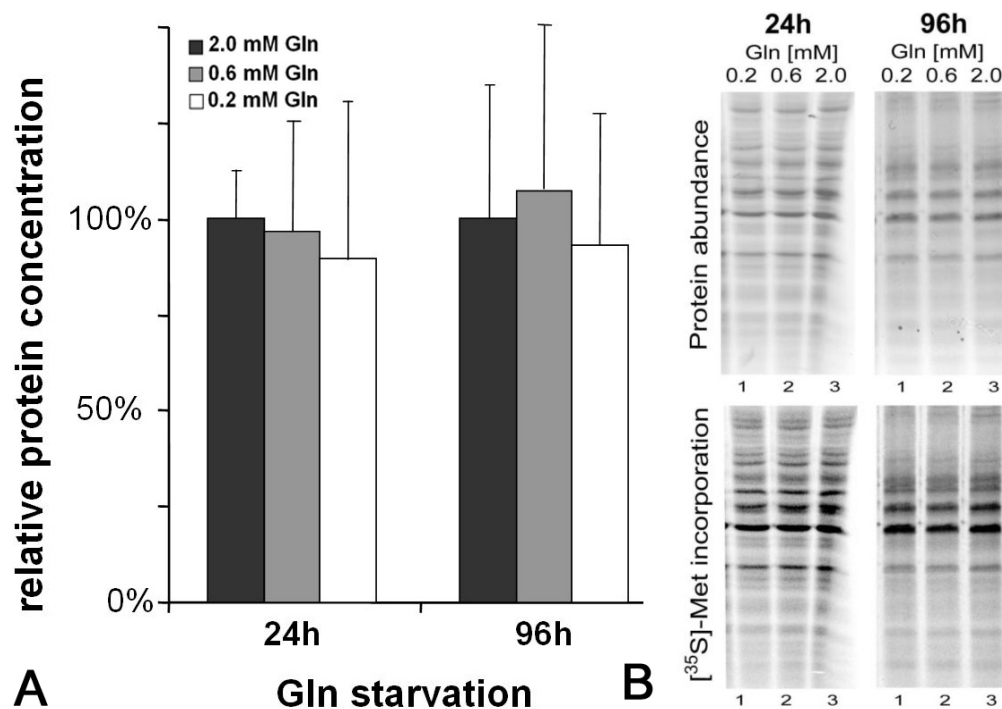


Figure 4. Cellular protein content and synthesis in Gln-depleted U937 cells. U937 cells were cultured in medium supplemented with various Gln concentrations for 24h or 96h. (A) Cellular protein content. At both time points, cells were harvested and free amino acids were removed by precipitation with 20% trichloroacetic acid at 0°C for 1h. After resolubilisation of the protein-pellet, protein content was determined by a Bradford assay. The mean protein content (n = 8) in control cells (2.0 mM Gln) was set to 100%. (B) Overall protein synthesis. Two hours before the end of the experiment [³⁵S]-Met was added to the cell suspension for incorporation into the synthesised proteins. At the end of the experiment, cells were harvested and a SDS polyacrylamide gel electrophoresis was performed. The abundance of the separated proteins was visualised by staining with ruthenium II tris (bathophenanthroline disulfonate) (upper two panels). The [³⁵S]-Met incorporation, which is a measure for protein synthesis, was visualised by autoradiography (lower two panels).

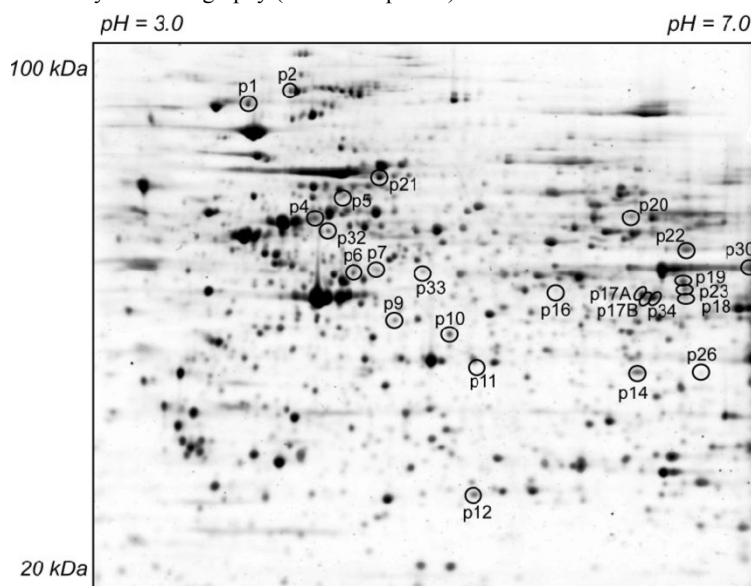


Figure 5. Two-dimensional gels of proteins from U937 cells. Cells were lysed and proteins were separated by two-dimensional gel electrophoresis. Gel were stained with ruthenium II tris (bathophenanthroline disulfonate). First dimension: pH range from 3 to 7. Second dimension from 20 kDa to 100 kDa. The circles indicate the position of the protein spots listed in tables 1 and 2.

Adaptation to glutamine starvation

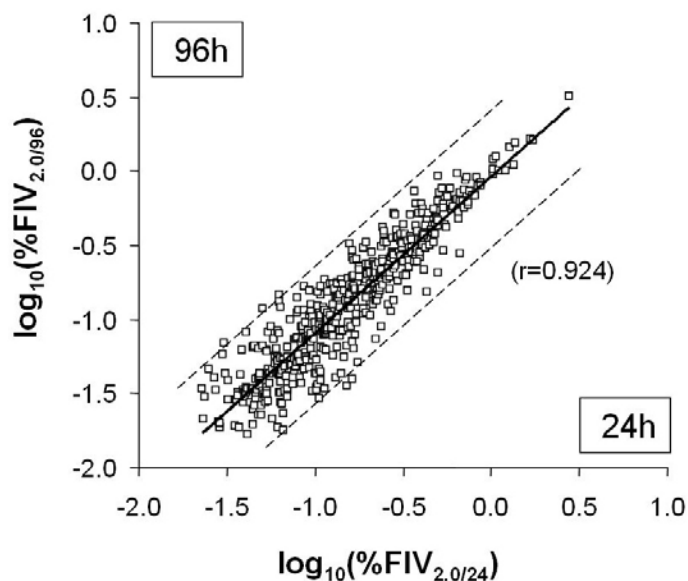


Figure 6. Scatter blot comparing the protein abundance in cells cultured for 24 h or 96 h at 2.0 mM Gln. Cells were harvested and proteins were separated by 2D gel electrophoresis. Protein spots were stained with ruthenium II tris (bathophenanthroline disulfonate) and the fluorescence intensity volume (FIV) of every spot was measured. Every cross in the diagram indicates the spot abundance (expressed as \log_{10} value of %FIV) in the 2.0/24 gel (X-axis) and in the 2.0/96 gel (Y-axis). The straight line represents the linear regression, the broken lines represents the 99% coincidence intervals. The correlation coefficient is indicated in brackets.

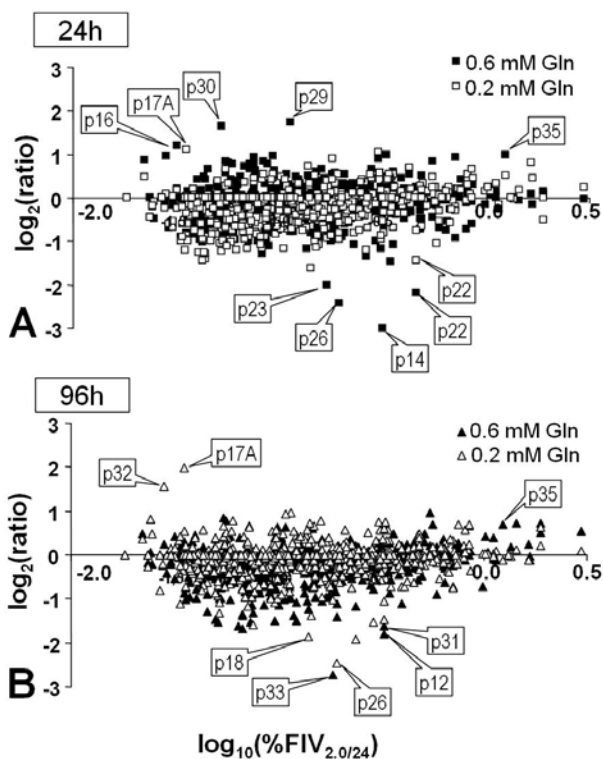


Figure 7. MA blots showing differences in the expression profiles of Gln depleted cells (0.6 mM Gln or 0.2 mM Gln) in comparison to cells cultured in complete medium (2.0 mM Gln). U937 cells were cultured in medium supplemented with 2.0 mM, 0.6 mM and 0.2 mM Gln for 24h (A) and for 96h (B). Then proteins were extracted and separated in a 2D gel electrophoresis. Protein spots were visualised by ruthenium II staining and the %FIV was used as a measure for abundance of every single spot. The MA blot shows on the Y-axis the $\log_2(\%FIV_{0.6 \text{ or } 0.2} / \%FIV_{2.0})$ and in the X-axis the $\log_{10}(\%FIV_{2.0/24})$ values. The labelled spots exhibit an disproportional high or low $\log_2(\text{ratio})$. Their position on the 2D gel is shown in Figure 6.

Adaptation to glutamine starvation

Table 1. Protein spots significantly increased or decreased by Gln-starvation

Spot ID	$\log_2(\text{Ratio to 2mM Gln})^1$			
	0.6mM Gln		0.2mM Gln	
	24h	96h	24h	96h
p1	-1,16	-0,08	-0,52	0,76
p2	-	-0,16	-0,64	-
p4	-1,09	-0,53	-0,8	-0,76
p14	-3	-0,32	-0,3	-0,15
P17A	-	0,29	1,11	1,99
P17B	-	0,18	0,17	-0,58
p18	-0,82	-1,24	-1,62	-1,87
p19	-1,36	-0,4	-1,19	-0,24
p20	-0,89	-0,66	-1,08	-0,27
p21	-0,4	0,21	-0,43	0,11
p22	-2,19	-0,31	-1,43	-0,63
p23	-2,02	-0,43	-0,93	-0,57
p16	1,2	-0,55	0,18	-0,04
p32	0,96	0,24	0,1	1,55
p34	0,69	-0,04	0,18	-0,16
p26	-2,44	-1,4	-0,82	-2,45
p30	1,66	-	-0,18	-0,24
p5	0,2	-1,39	-0,38	-
p6	-0,31	-1,53	-1,06	-0,92
p7	0,37	-1,05	0,28	-0,28
p11	0,44	-0,46	0,15	-0,1
p12	0,47	-1,79	0,08	-1,04
p33	-0,92	-1,69	-0,43	-0,42

¹ Values correspond to the logarithm on base 2 of the ratio between the abundance (%FIV) of protein spots from Gln starved cells (0.6 mM Gln or 0.2 mM Gln) with that of spots from control cells (2.0 mM Gln). The experiment was performed two times. The listed protein spots fulfil one of the following inclusion criteria: a $\log_2(\text{ratio})$ higher than +1 or lower than -1 in both experiments, or a $\log(\text{ratio})$ higher than 0.4 or lower than -0.4 which was nearly identical in both experiments (max. error 10 %), or absence of the spot in the Gln-starvation group in both experiments. The values which fulfil these conditions are indicated in bold. Spots are sorted according to their expression levels as described in the discussion section.

Table 2. Protein identification results

Spot ID	SwissProt Acc. Nr.	Protein Name	Spot position		Theoretical ¹		Sequence Coverage ²	MOWSEScore PMF ³	Confirmation by PSD or CID MS/MS
			MW [kDa]	pI	MW [kDa]	pI			
			100	5.2	94	5.2			
p4	Q12874	Splicing factor 3A subunit 3 (SAP 61)	60	5.3	59	5.3	17 %	63	PSD
p5	-	n.i.							
p6	P52597	Heterogeneous nuclear ribonucleoprotein F (hnRNP F)	45	5.5	46	5.4	26 %	125	PSD
p7	-	n.i.							
p11	Q96C86	Scavenger mRNA decapping enzyme (DcpS)	35	6.0	39	5.8	16 %	83	CID
p12	-	n.i.							
p14	P40925	Malate dehydrogenase (MDH)	35	6.5	36	6.9	40 %	170	PSD
p16	-	n.i.							
p17A	P49411	Elongation factor Tu ₁ (EF-Tu)	42	6.5	50	7.3	30 %	139	CID
p17B	P49411	Elongation factor Tu ₂ (EF-Tu)	42	6.5	50	7.3	33 %	117	PSD
p18	-	n.i.							
p19	P52209	6-phosphogluconate dehydrogenase (PGDH)	45	6.8	53	6.9	20 %	113	PSD
p20	-	n.i.							
p21	P08133	Annexin 6 (A6)	70	5.5	76	5.4	26 %	155	PSD
p22	P00367	Glutamate dehydrogenase 1 (GDH)	55	6.6	61	7.7	23 %	90	PSD
	P25705	ATP Synthase α -chain (ATP-S α)	55	6.6	60	9.2	19 %	86	PSD
p23	P62195	26S protease regulatory subunit 8 (p45/SUG)	42	6.6	46	7.1	31 %	156	---
p26	-	n.i.							
p30	-	n.i.							
p32	-	n.i.							
p33	-	n.i.							
p34	P22234	Phospho ribosyl amino imidazole-succino carboxamide synthase (PUR6)	42	6.5	47	7.1	22 %	88	CID

¹ Calculated from sequences retrieved from SwissProt database using expasy "Compute pI/Mtool", http://www.expasy.org/tools/pi_tool.html, ² Calculated as number of amino acids spanned by found peptides divided by sequence length of mature protein, ³ Probability based MOWSE Score for the PMF search result as calculated by MASCOT algorithm. Scores > 60 reflect significant results

Adaptation to glutamine starvation

analysis on the result. In contrast, some protein spots show disproportionately high or low ratios (e.g. in Figure 7A, 0.6 mM Gln: spots p16, p30, p29 and p35 are increased and spots p23, p26, p14, and p22 are decreased). This result showed that although the expression of most proteins was unaffected by Gln-starvation the expression of some proteins was specifically increased or decreased. The list of protein spots which expression was significantly influenced by Gln-starvation is shown in table 1. The quantitative evaluation indicated that different protein species reacted highly specific. About 5 protein spots were increased within 24h at 0.6 mM Gln whereas 7 were decreased (table 1, second column). Three were even non-detectable in 0.6/24 gels (p2, p17A, p17B). The protein p2 for example was nicely detectable in 2.0/24 control gels but absent in the 0.6/24 gels. Then it returned to normal levels within 96 h after start of the experiment (third column). At 0.2 mM Gln, however, p2 behaved completely different: after 24 h the protein spot was slightly reduced but disappeared at 96 h. To elucidate the identity of the proteins, spots were cut out, the isolated protein was enzymatically cleaved by trypsin and the resulting peptides were analysed by PMF applying either vacuum or AP-MALDI mass spectrometry. Final identity was confirmed by PSD or low energy CID experiments generating unambiguous sequence tags. The results are shown in table 2. The spots p17A and p17B contained both the same protein (EF-Tu). In contrast, spot p22 contained GDH as well as ATP synthase alpha-chain. The MOWSE score for both proteins was above 80 and the score for a mixture was 173, which clearly confirmed the presence of the two proteins. Some proteins showed a considerable difference between the theoretical pI and the pI estimated from the spot position on the 2D gel (e.g. p22, p23, p34). It can be assumed that this was due to post-translational modifications. Unfortunately, some low abundant proteins remained non-identifiable partly due to the applied wide IPG strips. The list includes metabolic enzymes (MDH, PGDH, GDH), proteins involved in RNA processing and degradation (hnRNP F and DcpS), proteins involved in protein synthesis and degradation (EF-Tu and p45/SUG) as well as a stress protein (Hsp-4). These results show that although long-term Gln-starvation has no effect on the overall protein synthesis rate it leads to specific changes in the protein expression profile.

5. DISCUSSION

The present study reveals that Gln-utilising monocytic U937 cells activate specific adaptive mechanisms in response to Gln-starvation. The U937 cell line was derived from malignant cells obtained from the pleural effusion of a patient with histiocytic lymphoma (20). These cells are well known to take up Gln (4, 21) and to utilise it for ATP synthesis (1). U937 cells react in response to Gln-starvation similarly as monocytes in respect to their overall protein synthesis and breakdown (1), their cellular hydration state (5, 6, 22) and their Hsp70 expression (4, 23). However, they are proliferating cells and differ considerably from primary monocytes. Thus, U937 cells are a well suited model system for Gln-utilising cells, however, it cannot be concluded that primary

monocytes would behave identically. U937 cells show an optimal growth at 2.0 mM Gln in the medium. Lower Gln concentrations lead to a decrease in DNA synthesis rate and to a shift in the cell cycle from S-phase to G1/G0-phase (22). Thus, the optimal Gln concentration for U937 cells is far above the concentration in human blood from which these cells originally derive (plasma Gln concentration is about 0.6 mM). This might be due to the well known higher Gln utilisation rate and glutaminolysis in tumor cells (24, 25). Therefore, we regarded 2 mM Gln in the medium as normal control condition.

To expose U937 cells to Gln-starvation we cultured them in a medium with 0.6 mM or 0.2 mM Gln. The first one corresponds to the Gln concentration in plasma and was considered a mild Gln-starvation. The latter one is clearly below all ever measured pathological plasma concentrations in critically ill patients (around 0.4 mM Gln) and is therefore considered a strong Gln-starvation. Both conditions, mild and strong Gln-starvation, lead to a decrease of Gln_i below 10 % of control within the first day. A similar drop of Gln_i was recently found to occur in U937 cells already within few hours after complete withdrawal of Gln from the medium (4). This decrease was accompanied by a reduction of intracellular free glutamate. Thus, U937 cells cultured under normal conditions seems to have a very high Gln utilisation, which leads to an immediate drop of Gln_i when the extracellular source of Gln is not available anymore. However, cells cultured at 0.6 mM Gln showed in the present study an increase of Gln_i in the second 24 h to about 40 % of control. This indicates that during times of a reduced Gln supply, cells activate adaptive mechanisms in order to increase their intracellular Gln. However, these mechanisms are only able to compensate for a mild Gln-starvation. The incomplete recovery of Gln_i at 0.6 mM Gln in the medium (2,414 nmol/ μ g protein instead of 5,850 nmol/ μ g protein at 2.0 mM Gln) is accompanied by a complete restoration of the normal proliferation rate and of the normal ATP levels within four days (see Figure 1). This suggests that either a Gln_i level of about 2,500 nmol/ μ g protein is sufficient for these cellular functions or that the cells find a way to compensate for the still missing 3,300 nmol/ μ g protein Gln_i .

Cellular ATP level is tightly controlled and the AMP-activated kinase (AMPK) plays a central role in this regulation (26). Whenever the cellular ATP:ADP ratio falls, owing to a starvation that inhibits ATP production or to an increased ATP consumption, this is amplified by adenylate kinase into a much larger increase in the AMP:ATP ratio. AMP activates the AMPK by binding and phosphorylation. The present study shows that AMPK is activated by the 20 % decrease of ATP in response to a mild Gln-starvation. This activation disappears within the following three days, when the ATP level returned to normal. However, ATP levels remained stably decreased by 40 % and AMPK remained stably activated throughout the experiment in cells exposed to strong Gln-starvation. Activation of AMPK downregulates biosynthetic pathways such as fatty acid and cholesterol biosynthesis, yet switches on catabolic pathways that generate ATP, such as fatty acid oxidation,

Adaptation to glutamine starvation

glucose uptake and glycolysis (an interaction of AMPK with glutaminolysis is hitherto unknown). Therefore, we propose that AMPK contributes to the restoration of normal ATP levels after four days of mild Gln-starvation. However, we cannot estimate whether AMPK is involved in the regain of Gln_i under these conditions. In a recent publication, we could show that U937 cells downregulate the expression of the cell protective protein Hsp70 immediately in response to Gln-starvation by a reduced mRNA stability (4). This reduction is supposed to be responsible for the increased susceptibility of Gln-starving U937 cells to apoptosis (2). The present study revealed that Hsp70 expression decreased within 24 h in response to 0.6 mM and 0.2 mM Gln. After four days of Gln-starvation Hsp70 expression returned to normal at 0.6 mM Gln but remained decreased at 0.2 mM Gln. Thus, Hsp70 expression recovers at mild Gln-starvation and shows a similar dynamic as Gln_i, proliferation rate, ATP, and AMPK.

All these results show that U937 cells activate an adaptive mechanism in response of Gln-starvation, which enables them to overcome the lack of Gln if there is 0.6 mM Gln still available. In order to characterise the proteins involved in the adaptive response to Gln-starvation we performed a proteomic study. Cells were cultured under mild and strong Gln-starvation for 24 h and 96 h and the expression level of every single protein was compared to its expression level in control cells. This resulted in a subdivision of all proteins in four groups: (I) transiently changed proteins, which return to normal expression after 96 h, (II) permanently changed proteins, (III) posterior changed proteins which show normal expression at 24 h but changed expression at 96 h and finally (IV) unchanged proteins. The great majority of proteins belonged to the unchanged group. Only 23 proteins were found to be affected by Gln-starvation (listed in table 1). However, it was not possible to reveal the identity of all of them. Twelve proteins were transiently down-regulated at 0.6 mM Gln: p1, p2 (Hsp-4), p14 (MDH), p17 (EF-Tu), p18, p19 (PGDH), p20, p21 (A6), p22 (GDH and ATP-S alpha), and p23 (p45/SUG). The protein spot p22 contained GDH as well as ATP-S alpha and it was not clear whether both proteins were affected similarly by Gln-starvation. Hsp-4 is a member of the Hsp70 family and exhibited a similar dynamic in its expression as Hsp70: at 0.6 mM Gln, Hsp-4 expression decreased within 24 h but regains normal values within 96 h. In contrast, at 0.2 mM Gln it decreased within 24 h and disappeared completely after 96 h. This indicates that the inhibitory effect of Gln-starvation on Hsp70 expression shown in our previous study (4) affects also other members of the heat shock protein family. However, it remains unclear whether Hsp-4 is also down-regulated by a decreased mRNA stability. Glutamate dehydrogenase (GDH) and malate dehydrogenase (MDH) are involved in glutaminolysis (27). A down-regulation of these enzymes in times of reduced availability of Gln has a high probability. The fact that this reduction was transiently at 0.6 mM Gln indicates that the regained Gln_i level in the later phase of the experiment was sufficient to reactivate glutaminolysis. However, a more detailed metabolic analysis would be necessary to prove this assumption. The

enzyme 6-phosphogluconate dehydrogenase (PGDH) catalyses the third step in the pentose phosphate pathway, which is on one hand a bypass pathway for glycolysis and on the other hand provides substrate for nucleotide synthesis (28, 29). The reason for down-regulation of this enzyme remains unclear. Three proteins are transiently up-regulated by Gln-starvation: p16, p32 and p34 (PUR6). Phosphoribosyl aminoimidazole-succinocarboxamide synthase (PUR6) catalyses an important step in the synthesis of 5-aminoimidazole-4-carboxamide-1-beta-D-ribofuranoside (AICAR) which is a precursor of IMP and finally all *de novo* synthesised nucleotides. AICAR is a well known strong and specific activator of AMPK (30) and has a variety of effects on cellular metabolism including increased fatty acid and glucose oxidation, enhanced GLUT2-dependent glucose uptake, suppressed IL-2 expression and stimulated IL-6 production. Two proteins were permanently changed by Gln-starvation: p26 was decreased and p30 was increased. However, it was not possible to identify these proteins due to their low abundance. Six proteins were posterior down-regulated by Gln-starvation: p5, p6 (hnRNP-F), p7, p11 (DcpS), p12, and p33. Both, the heterogeneous nuclear ribonucleoprotein F (hnRNP-F) and the scavenger mRNA decapping protein (DcpS), are involved in mRNA metabolism: hnRNP-F in mRNA processing and DcpS in mRNA degradation. The down-regulation of these proteins in the late phase of adaptation may suggest that these proteins are needed to a lesser extent in times of reduced Gln supply.

The present study revealed that Gln-utilising U937 cells activate adaptive mechanisms in response to Gln-starvation, which enable them to overcome a Gln shortage. However, when the level of available Gln falls below a critical concentration, these adaptive mechanisms are not sufficient to counteract the lack of the amino acid. We could identify a list of proteins, which are involved in this process. Although this list is incomplete and the role of all proteins is not yet understood, it gives an idea of the complex processes during Gln-starvation and might be the starting point for many future studies on Gln-starvation.

6. ACKNOWLEDGMENT

Maja Munk Eliassen, and Wolfgang Winkler contributed equally to the manuscript. This work was supported by the European Society of Parenteral and Enteral Nutrition (ESPEN Fellowship 2001). We thank Susanne Oehler for carefully reading the manuscript.

7. REFERENCES

1. Zellner, M., C. Gerner, M. Eliassen, S. Wurm, J. Pollheimer, A. Spittler, C. Brostjan, E. Roth & R. Oehler: Glutamine starvation of monocytes inhibits the ubiquitin-proteasome proteolytic pathway. *BBA - Molecular Basis of Disease*, 1638, 138-148 (2003)
2. Exner, R., G. Weingartmann, A. Spittler, M. Brabec, E. Roth & R. Oehler: Glutamine deficiency renders human monocytic cells more susceptible to specific apoptosis triggers. *Surgery*, 131, 75-80 (2002)

Adaptation to glutamine starvation

3. Fumarola, C., A. Zerbinì & G. G. Guidotti: Glutamine deprivation-mediated cell shrinkage induces ligand-independent CD95 receptor signaling and apoptosis. *Cell Death Differ*, 8, 1004-13 (2001)
4. Eliassen, M., M. Brabec, C. Gerner, J. Pollheimer, H. Auer, M. Zellner, G. Weingartmann, F. Garo, R. Roth & R. Oehler: Reduced stress tolerance of glutamine-deprived human monocytic cells is associated with selective down-regulation of Hsp70 by decreased mRNA stability. *Journal of Molecular Medicine*, 84, 1-12 (2005)
5. Spittler, A., S. Holzer, R. Oehler, G. Boltz-Nitulescu & E. Roth: A glutamine deficiency impairs the function of cultured human monocytes. *Clin-Nutr.*, 16, 97-99 (1997)
6. Spittler, A., S. Winkler, P. Göttinger, R. Oehler, M. Willheim, C. Tempfer, G. Weigel, R. Függer, G. Boltz Nitulescu & E. Roth: Influence of glutamine on the phenotype and function of human monocytes. *Blood*, 86, 1564-9 (1995)
7. Oehler, R., B. Schmierer, M. Zellner, R. Prohaska & E. Roth: Endothelial Cells Downregulate Expression of the 70 kDa Heat Shock Protein during Hypoxia. *Biochem-Biophys-Res-Commun.*, 274, 542-547 (2000)
8. Luo, J. L., F. Hammarqvist, I. A. Cotgreave, C. Lind, K. Andersson & J. Wernerman: Determination of intracellular glutathione in human skeletal muscle by reversed-phase high-performance liquid chromatography. *J Chromatogr B Biomed Appl*, 670, 29-36 (1995)
9. Zellner, M., W. Winkler, H. Hayden, M. Diestinger, M. Eliassen, B. Gesslbauer, I. Miller, M. Chang, A. Kungl, E. Roth & R. Oehler: Quantitative Validation of different Protein Precipitation Methods in Proteome Analysis of Blood Platelets. *Electrophoresis*, 26, 2481-2489 (2005)
10. Laemmli, U. K.: Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, 227, 680-5 (1970)
11. Lamanda, A., A. Zahn, D. Roder & H. Langen: Improved Ruthenium II tris (bathophenanthroline disulfonate) staining and destaining protocol for a better signal-to-background ratio and improved baseline resolution. *Proteomics*, 4, 599-608 (2004)
12. Jensen, O. N., M. Wilm, A. Shevchenko & M. Mann: Sample preparation methods for mass spectrometric peptide mapping directly from 2-DE gels. *Methods Mol Biol*, 112, 513-30 (1999)
13. Pluskal, M. G.: Microscale sample preparation. *Nat Biotechnol*, 18, 104-5 (2000)
14. Vorm, O., P. Roepstorff & M. Mann: Improved Resolution and very high sensitivity in MALDI TOF of matrix surfaces made by fast evaporation. *Analytical Chemistry*, 66, (1994)
15. Clauser, K. R., P. Baker & A. L. Burlingame: Role of accurate mass measurement (± 10 ppm) in protein identification strategies employing MS or MS/MS and database searching. *Anal Chem*, 71, 2871-82 (1999)
16. Perkins, D. N., D. J. Pappin, D. M. Creasy & J. S. Cottrell: Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis*, 20, 3551-67 (1999)
17. Bairoch, A. & R. Apweiler: The SWISS-PROT protein sequence database: its relevance to human molecular medical research. *J Mol Med*, 75, 312-6 (1997)
18. Wilkins, M., K. Williams, R. Appel & D. Hochstrasser: Proteome Research: New Frontiers in functional Genomics. Principles and Practice (ed 1). Springer Verlag, Berlin Heidelberg New York (1997)
19. Karas, M. & F. Hillenkamp: Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Anal Chem*, 60, 2299-301 (1988)
20. Sundstrom, C. & K. Nilsson: Establishment and characterization of a human histiocytic lymphoma cell line (U-937). *Int. J. Cancer*, 17, 565-77 (1976)
21. Haas, H. S., R. Pfragner, V. Siegl, E. Ingolic, E. Heintz & K. Schauenstein: Glutamate receptor-mediated effects on growth and morphology of human histiocytic lymphoma cells. *Int J Oncol*, 27, 867-74 (2005)
22. Spittler, A., R. Oehler, P. Goetzinger, S. Holzer, C. M. Reissner, F. Leutmezer, V. Rath, F. Wrba, R. Függer, G. Boltz Nitulescu & E. Roth: Low glutamine concentrations induce phenotypical and functional differentiation of U937 myelomonocytic cells. *J Nutr*, 127, 2151-7 (1997)
23. Pollheimer, J., M. Zellner, M. Eliassen, E. Roth & R. Oehler: Increased susceptibility of Glutamine-depleted monocytes to fever-range hyperthermia: The role OF 70 kDa Heat shock protein. *Annals of Surgery*, 241, 349-55 (2005)
24. Newsholme, E. A. & M. Board: Application of metabolic-control logic to fuel utilization and its significance in tumor cells. *Adv Enzyme Regul*, 31, 225-46 (1991)
25. Mazurek, S. & E. Eigenbrodt: The tumor metabolome. *Anticancer Res*, 23, 1149-54 (2003)
26. Hardie, D. G.: The AMP-activated protein kinase pathway--new players upstream and downstream. *J Cell Sci*, 117, 5479-87 (2004)
27. Newsholme, E. A., B. Crabtree & M. S. Ardawi: The role of high rates of glycolysis and glutamine utilization in rapidly dividing cells. *Biosci Rep*, 5, 393-400 (1985)

Adaptation to glutamine starvation

28. Ahmed, N., J. F. Williams & M. J. Weidemann: Glycolytic, glutaminolytic and pentose-phosphate pathways in promyelocytic HL60 and DMSO-differentiated HL60 cells. *Biochem Mol Biol Int*, 29, 1055-67 (1993)

29. Rippa, M., P. P. Giovannini, M. P. Barrett, F. Dallochio & S. Hanau: 6-Phosphogluconate dehydrogenase: the mechanism of action investigated by a comparison of the enzyme from different species. *Biochim Biophys Acta*, 1429, 83-92 (1998)

30. Corton, J. M., J. G. Gillespie, S. A. Hawley & D. G. Hardie: 5-aminoimidazole-4-carboxamide ribonucleoside. A specific method for activating AMP-activated protein kinase in intact cells? *Eur J Biochem*, 229, 558-65 (1995)

Abbreviations: 2DE: Two-dimensional gel electrophoresis; A6: Annexin 6; AMPK: AMP activated kinase; AP-MALDI: Atmospheric pressure-MALDI; ATP: Adenosyl triphosphate; ATP-Sa: ATP Synthase a-chain; CID: Collision induced dissociation; DcpS: Scavenger mRNA decapping enzyme; DTT: Dithiotreitol; EF-Tu: Elongation factor Tu; FCS: Fetal calf serum; FIV: Fluorescence intensity volume; GDH: Glutamate dehydrogenase 1; Gln: Glutamine; GSH: Glutathione; hnRNP F: Heterogeneous nuclear ribonucleoprotein F; Hsp-4: Heat shock 70 kDa protein 4; Hsp70: 70kDa heat shock protein; IEF: Isoelectric focusing; MALDI: Matrix-Assisted Laser Desorption/Ionization; MDH: Malate dehydrogenase; MS: Mass spectrometry; OPA: o-Phthalaldehyde; p45/SUG: 26S protease regulatory subunit 8; PBS: Phosphate buffered saline; PDF: Pulsed dynamic focusing; PGDH: 6-Phosphogluconate dehydrogenase; PUR6 : Phospho ribosyl amino imidazole-succino carboxamide synthase; SAP 61: Splicing factor 3A subunit 3; SDS-PAGE: Sodium dodecyl sulfate polyacryl amide gel electrophoresis; TCA: Tricloacetic acid; TNF-alpha: Tumor necrosis factor alpha; TOF: Time of flight

Key Words: Glutamine, Amino acid starvation, Starvation response, Stress proteins, Proteomics

Send correspondence to: Dr Rudolf Oehler, Department of Surgery, Medical University of Vienna, AKH (8G9.05), Waehringer-Guertel 18-20, A-1090 Vienna, Austria, Tel: 43-1-40400-6979, Fax: 43-1-40400-6782, E-mail: rudolf.oehler@meduniwien.ac.at

<http://www.bioscience.org/current/vol11.htm>

3 Biological variation of platelet proteome in the elderly population and its implication for biomarker research

Manuscript

**Biological variation of platelet proteome in the elderly population
and its implication for biomarker research**

Wolfgang Winkler*¹, Maria Zellner*², Michael Diestinger², Daniel Halbauer²,
Alexandra Goll³, Sonja Zehetmayer³, Peter Bauer³, Eduard Rappold⁴, Martina
Marchetti¹, Erich Roth², Günter Allmaier¹ and Rudolf Oehler²

Author affiliation:

¹Institute of Chemical Technologies and Analytics, Vienna University of Technology,

²Department of Surgery and ³Section of Medical Statistics, Medical University of

Vienna and ⁴Geriatric centre of the Otto Wagner hospital, Vienna, Austria.

*Both authors contributed equally to the manuscript

Corresponding author:

Rudolf Oehler, Department of Surgery, Medical University of Vienna, AKH (8G9.05),

Waehringerguertel 18-20, A-1090 Vienna, Austria.

E-mail: rudolf.oehler@meduniwien.ac.at;

Tel.: +43-1-40400-6979; Fax: +43-1-40400-6782

Running title:

Biological variation of the platelet proteome

Abstract:

Knowledge about the extent of both, the technical variation of the method and the total variation experienced between individual samples is of great importance for the design of not only proteomic, but every clinical study. The smallest signal difference, detectable with statistic significance depends on the technical variation. The total variation between individuals, together with the sample size, defines the smallest statistically significant detectable signal difference when comparing two or more groups of individuals. We used platelet proteins isolated from 20 healthy human volunteers aged from 56-100 years to determine the technical and total variation experienced in a proteome analysis using 2D-DIGE to investigate the pI ranges 4-7 and 6-9. Our results show a technical variability of about 5% and a total variability of 16%. With these data, it is possible to give an estimation for the sample size needed to detect given protein expression differences with statistic significance. Additionally, we provide a list of proteins which show very low total variance and therefore can be used for standardisation and normalisation.

Introduction:

Due to the ability of the proteomics technology to separate and simultaneously quantify thousands of proteins, the number of studies to search for biomarkers on protein level has dramatically increased during the last years. Biomarkers are used to differentiate between two or more different biological states and to monitor disease progress or the success of medical treatment [1]. To fulfill these tasks, there are not only strict requirements to each biomarker-candidate in terms of selectivity and specificity, but also to the precision of the measurement. So the knowledge of the extent of variation between the individuals within each group is crucial for the study design as well as the selection of the method of measurement. The total variation experienced between individuals within one group is the result of the biological variation caused by factors like sex, age, genetic background or lifestyle and the technical variation introduced by sample handling and the method of measurement itself. This study investigates the extent of variation every search for biomarkers in human blood platelets has to deal with.

Platelets are the smallest cells circulating in human bloodstream and are easy to obtain in sufficient amount and purity. They are responsible for the maintenance of vascular integrity and are also involved in inflammation and wound healing [2]. Platelets are anucleate cytoplasmatic fragments released from megacaryocytes in the bone marrow. During the platelet biogenesis, organelles, especially the mitochondria and the α - and dense granulae are actively transported into the platelets via the channel-like proplatelets [3]. Furthermore, platelets received a certain set of mRNAs during their biogenesis from the megacaryocytes and are still capable of protein synthesis and processing [4].

Epidemiological studies show a dramatic increase in the rates of venous thrombotic events like myocardial infarction and stroke in the population with advancing age [5]. These findings obviously are the result of an age-related disorder in the whole haemostatic system including the platelets. Alterations in protein expression levels which cause an imbalance between pro-coagulant system and anti-coagulant and fibrinolytic system as well as alterations in platelet function resulting in a decreased activation threshold are described [6].

In this study we used two-dimensional gel electrophoresis to separate the platelet proteins by their isoelectric point and molecular weight.

We used twodimensional differential gel electrophoresis [7] and all samples were separated together with an internal standard which was created by combining the same amount of protein from each sample. Both, sample and internal standard were labelled with different fluorescence dyes prior to electrophoresis and therefore can be detected with high sensitivity by laser imaging. By normalising the quantitation information of every spot by the corresponding spot in the internal standard, and using only these ratios to compare quantitation information across gels, 2D-differential gel electrophoresis (2D-DIGE) is significantly less susceptible to errors in quantification caused by suboptimally reproduced 2D spot patterns.

Several proteomics studies have focused on different aspects in platelet research using 2D electrophoresis to separate the platelet proteins and subsequent mass spectrometry to identify them: 2D maps and the total proteome [8], [9], the platelet secretome [10], the processes during platelet activation via analysis of phosphoproteins [11], [12]. But until now, no quantitative proteomic studies which aim to find differences in protein expression levels between two or more groups of individuals have been published.

Here we use the quantitation information obtained from 2D-DIGE to determine the extent of variation in protein expression levels experienced between the individuals of a group of 20 healthy volunteers. The proteins quantified were identified by peptide mass fingerprinting (PMF) and MALDI-MS/MS experiments. This group consists of 14 women and 6 men with an average age of 77 years (age spanned from 56 to 100 years) and is due to its heterogeneous composition best suited to be an universally applicable control group for many different aspects in platelet research. Our findings will be advantageous for the design of future biomarker searches not only in platelets, but can also serve as guiding values for other studies using different tissue types.

Experimental Procedures:

Blood sampling

With approval of the local ethics committee, blood was drawn from 20 healthy fasting volunteers (6 male, 14 female with average age of 77.3 ± 12 years). Blood was drawn without stasis from the antecubital vein into vacuum tubes (Greiner Bio-One GmbH, Kremsmünster, Austria) containing 0.129 mol/L sodium citrate as anticoagulant (mixing ratio 1:9 with blood). The first drawn tube was discarded.

2D sample preparation

The platelet isolation as well as the protein precipitation was carried out as described previously [13]. Briefly, centrifugation of the blood samples followed by size exclusion chromatography of the supernatant (platelet rich plasma) was used to obtain a pure platelet suspension in PBS. Out of this suspension, the platelet proteins were precipitated using trichloroacetic acid, washed and stored as pellets at -70°C in acetone (p.a., Merck, Darmstadt, Germany) enriched with 20 mM DTT (Roche Diagnostics, Mannheim, Germany). Before usage, the protein pellets were washed with ice-cold acetone to remove DTT and dried by evacuation. Subsequently, the pellets were resolubilised over night at 4°C using 70 μL of 2D sample buffer (7 M urea, 2 M thiourea, 4% CHAPS, 30 mM TRIS-HCl; pH=8.5) per 100 million platelets. The protein concentration of the samples was determined in triplicate using a CBB based assay kit (Coomassie Plus, Pierce, Rockford, IL, USA) via a BSA standard curve. The internal standard was composed by pooling the same amount of protein of every sample. Internal standard and samples were stored at -70°C as aliquots of appropriate size. Labelling with fluorescent cyanine dyes (CyDyes, GE Healthcare, Giles, UK) was carried out according to the manufacturer's instructions. However, the ratio of CyDye to protein was slightly reduced to 7 pmol of dye per μg protein. The internal standard always was labelled using Cy2, the samples were labelled alternately with Cy3 and Cy5, respectively.

2D-Electrophoresis

Twenty-four cm pH 4-7 IPG-Drystrips (GE Healthcare, Giles, UK) were passively rehydrated for 14 hours at room temperature within a mixture containing 33 μg of labelled sample, 33 μg of labelled internal standard, 70 mM DTT, 0.5% ampholytes pH 4-7 (Serva, Heidelberg, Germany) and the volume of rehydration solution (7 M

urea, 2 M thiourea, 4% CHAPS) was added up to a final volume of 450 μ L. Isoelectric focussing was carried out on an IPGPhor device (GE Healthcare, Giles, UK) by applying 30 kVh. Twenty-four cm pH 6-9 IPG-Drystrips were passively rehydrated for 14 hours at room temperature within 450 μ L of rehydration solution containing 150 mM DTT and 2% ampholytes pH 6-9 (Serva, Heidelberg, Germany). A mixture containing 25 μ g labelled sample, 25 μ g labelled internal standard, 150 mM DTT, 2% ampholytes pH 6-9 and the volume of 2D sample buffer to reach 55 μ L was applied to the DryStrip via cup loading and IEF was performed by applying 30 kVh. Prior to SDS-PAGE the IPG-strips were equilibrated for 20 minutes within equilibration solution (50 mM TRIS-HCl, 6 M urea, 30% glycerol, 2% SDS, pH=8.8) containing 10 mg/mL DTT followed by equilibration for 15 minutes within equilibration solution containing 25 mg/mL Iodacetamide (Sigma, St. Louis, MO, USA). Thereafter the SDS-PAGE was performed on self cast isocratic gels (12.5% acrylamide) using the Ettan DALT six equipment (GE Healthcare Giles, UK). Electrophoresis conditions were: 35 V for the first hour, followed by 50 V for the next 1.5 hours and finally 110 V over night (16.5 hours) at 10°C.

Image acquisition and gel handling

The gels were scanned using a Typhoon TRIO scanner (GE Healthcare Giles, UK) with a resolution of 100 μ m. Photomultiplier voltage was adjusted before scanning to values where no channel had saturated spots (Actin-spot in pI-range 4-7 was always saturated) in pre-scan (1000 μ m resolution). After scanning, the gels were silver stained [14] and stored within 1% acetic acid at 4°C until protein identification.

Gel-Image processing and spot matching

Prior to spot detection, the gel images were edited using Image Quant software (version 5.2, GE Healthcare) to remove gel borders. Spot detection was performed on the edited images using the DeCyder module DIA (version 6.00.28) setting the target spot number to 2800 (pH range 4-7) and 2500 (pH range 6-9), respectively. From each gel spots which showed saturation (peak height above 65000) were excluded from further analysis. Next an exclusion filter for spots which show a slope greater than 1.4 or a volume below 25000 was applied to remove all spikes caused by dust particles as well as spots which were outlined in noise-only regions. Finally, all spots which have been removed by exclusion filters were manually reviewed to ensure no false exclusion. The fluorescent signal of a spot was background-corrected

and normalised by the DeCyder software (version 6.00.28) according to the manual: Briefly, for background correction the 10th percentile pixel value on the spot border was subtracted. For intra-gel normalisation a histogram was plotted for spot frequency against \log_{10} spot volume, a Gaussian distribution was fitted and the center of this curve was used as a correction factor for each spot. The resulting normalised volume was denoted as $Vnorm_g$. Each gel was added to the appropriate workspace and group and matched against the master gel using the DeCyder module BVA (version 6.01.02). Before the matching process was started the first time, up to 250 landmarks have been defined all over the gel with high density within regions where the algorithm used to fail correct matching. Subsequently, the matching process was started followed by reviewing all matches which were classified by the algorithm as level 1 match (automatic matches with high confidence) and confirming all correct matches. The cycle of matching, reviewing and confirming the matches was repeated until no new level 1 matches were found. Using the DeCyder module XMLToolbox, the spot raw data values especially the normalised volume $Vnorm_g$ for each gel were exported from DeCyder and analysed using a spreadsheet software.

Protein identification

The spots of interest were excised manually and subjected to *in-gel* digestion [15] using trypsin (from bovine pancreas, modified; sequencing grade, Roche, Mannheim, Germany). The digests were desalted and concentrated using ZipTips (C18, Millipore, Bedford, MA, USA) and analysed on a MALDI TOF/RTOF instrument (TOF2 [16], Shimadzu Biotech, Manchester, UK) applying thin layer preparation technique [17] using α -cyano-4-hydroxycinnamic acid (C2020, Sigma-Aldrich, St. Louis, MO, USA) as matrix. All mass spectra were recorded in positive ion reflectron mode with delayed extraction by accumulating up to 500 single unselected laser shots. Internal calibration was performed using the singly charged peptides from tryptic auto digestion at m/z 659.38, 805.42, 1153.57 and 2163.06 Da (monoisotopic masses). The lists of m/z -values for PMF searches were manually derived from unsmoothed mass spectra omitting those values related to tryptic auto digestion, contamination from keratins [18] and matrix cluster-ions [19]. The m/z -lists were submitted to the PMF search engines ProFound [20] and MASCOT [21] to search the SwissProt (version: 50.8) protein sequence database [22] (MASCOT only) and the non redundant database from NCBI (version 20061013) using settings for fixed

modification (carbamidomethylation), variable modification (oxidised Methionine and protein N-terminal acetylation), missed cleavages (1), species (homo sapiens), mass tolerance (± 0.15 Da for monoisotopic masses, ± 0.75 Da for average masses), molecular weight (value estimated from 2D gel $\pm 33\%$, ProFound only) and isoelectric point (value estimated from 2D gel ± 2 , ProFound only).

To confirm the PMF results, high energy CID experiments (collision gas: He, isolation width: precursor ± 5 Da, $E_{lab}=20$ keV) were performed on selected tryptic peptides using the same sample preparation technique as described above. Mass spectra were recorded in positive ion mode with delayed extraction by accumulating up to 3000 single unselected laser shots. The m/z-values for MS/MS ion search were manually derived as average masses from smoothed spectra (Savitsky-Golay algorithm [23]) and submitted to the MASCOT MS/MS search engine to search the SWISSPROT and NCBI nr database, using the same restrictions as described for PMF-searches except instrument type (MALDI-TOF-TOF) and mass tolerance which was set to ± 1 Da for product ions and ± 0.15 Da for the precursor ion.

A hit was considered to be significant if the scores obtained for PMF data and MS/MS data clearly exceeded each algorithm's significance threshold ($p < 0.05$).

Results:

Technical Variation

For quantification of biological variation between samples from different individuals it is essential to have an accurate detection system with a known technical variation. To determine the technical variation of each spot in a DIGE approach, we labelled the very same platelet protein extract with Cy2, Cy3 and Cy5. These extracts were mixed and separated on six different gels: three gels in the pI range of 4-7 and three gels in the pI range of 6-9. Typical gels are shown in figure 1. Then spots were detected and quantified on each gel. The fluorescence volume of each spot was background corrected and subjected to an intra-gel normalisation using the DeCyder6.0 software package algorithm. For inter-gel normalisation Cy3 and Cy5 signals were divided by their corresponding Cy2 signal on the same gel. This results in two "standardised abundance" (SA) values for each spot on each gel: one for Cy3 and one for Cy5. Since we analysed the same sample in all three fluorescence channel these ratios should be theoretically always one. Any observed deviation from the value one was defined as technical variation. The SA values were \log_2 -transformed to obtain a symmetrical distribution around zero. The $\log_2(\text{SA})$ values for each spot have been found to be normally distributed between the six different separations ($n = 2 \text{ dyes} * 3 \text{ gels} = 6$). A symmetrical normal distribution is a precondition for the statistical calculations used below. To quantify the technical variation the standard deviation (σ) of $\log_2(\text{SA})$ values was calculated for each spot. Figures 2 A and B show the spot frequency against σ . The great majority of spots have a very low variation: 95% of the spots in the acidic pI range have a σ below 0.185 and 95% of spots in the alkaline pI range have a σ below 0.224. The median σ is 0.0753 for acidic proteins and 0.0664 for alkaline proteins. A re-transformation from the \log_2 -space into the normal space revealed that this corresponds to a SA range from 0.941 to 1.048 (acidic proteins) and from 0.957 to 1.051 (alkaline proteins), respectively. Thus, the intensity of a typical spot of the very same sample varies between different DIGE separations by about $\pm 5\%$.

Variations between individuals

To evaluate the extent of variation between individuals, platelet protein extracts of 20 healthy volunteers were labelled either with Cy3 or Cy5 (see table 1). Every sample

was mixed with a Cy2 labelled internal standard which was generated by pooling an identical protein amount of all samples. These mixtures were separated on 40 2D gels: 20 in the pI range 4-7 and 20 in the pI range 6-9. Then spots were detected, quantified and SA values were calculated as described above. Only spots, which are present in the internal standard (Cy2) of at least 18 of 20 gels (90%), were included in the analysis (484 acidic protein spots and 424 alkaline protein spots). From each of these spots the average of $\log_2(\text{SA})$ and its standard deviation (σ) was calculated. Figures 2 C and D shows that the majority of σ values is much higher than in the experiment with identical samples described above. In addition, the σ values are distributed over a broad range. This indicates that every protein species has its own characteristic biological variation. Seven acidic and three basic spots have even a σ above 1. However 95% of acidic spots have a σ below 0.687 and 95% of alkaline spots have a σ below 0.581. The median σ is 0.242 and 0.246, respectively. This corresponds to a SA range from 0.79 to 1.15 and from 0.79 to 1.17. Thus, a typical platelet protein spot varies in abundance between different individuals by about $\pm 16\%$. The standard deviation shows neither a dependency on pI (figures 3 A and B), nor on molecular weight (figures 3 C and D), nor on protein abundance (figures 3 E and F).

The variations observed between individuals are clearly higher than the variations observed in the multiple analysis of a single sample described above. However, the distributions of σ values show a small overlap (figures 2 A,C and B,D). To estimate the proportion of the technical variation on the total variation measured between different individuals the two values were plotted against each other (figures 4 A and B). For the pI range of 4-7 the median σ between individuals is 0.242 (vertical line) while it is only 0.0752 (horizontal line) in the multiple analysis of the same sample. In the alkaline gel the median σ is 0.246 vs. 0.0664. Spots lying close to the 45 degree line showed a similar standard deviation in the analysis of different individuals as in the analysis of the same sample. However, almost all spots are below the 45 degree line. In addition, the very high standard deviation between individuals measured for some spots is not associated to a high technical variation. This indicates that for most spots the technical variation has only a minor contribution to the total variation measured in different biological samples.

A distinct group of spots show a very low biological variation: 20 acidic proteins (4.1% of all) and 22 alkaline proteins (5.2% of all) have a σ below 0.125. These proteins vary very little between individuals (below 9%) and might be therefore of special interest. We calculated the probability that a σ value is lower than 0.125 under the assumption that σ is normal distributed. Using the Bonferroni correction for multiple testing the probability is 0.000026 for acidic proteins ($n = 484$) and 0.00016 for alkaline proteins ($n = 424$). This strongly suggests that the low variation found for these proteins results from a biological effect and is not a consequence of statistical prerequisites. Eighteen clearly separated and sufficiently abundant spots of this group were selected for identification of the protein by PMF and high energy CID tandem MS (table 2). The positions of these spots on the 2D gel are indicated in figure 1. They are well distributed over the gels and include 11 cytoplasmatic, one mitochondrial, one ER lumen, one membrane, and two membrane associated proteins. Some of these proteins are involved in metabolism (fructose-bisphosphate aldolase A, GAPDH, triosephosphate isomerase, peptidyl-propyl cis isomerase). Others are involved in receptor binding and signalling (integrin alpha-IIb, guanine nucleotide-binding protein, 14-3-3 proteins gamma, theta and zeta/delta, growth factor receptor-bound protein 2), in cytoskeleton (myosin light polypeptide 6, actin-related protein and profilin-1) and in redox regulation (peroxiredoxin-6 and protein disulfide-isomerase). Because of their low biological variation these proteins might be suitable as reference proteins in future studies.

Implication for biomarker research

Many clinical proteomic studies aim to find biomarkers for a biological status of interest (e.g. a disease). Therefore proteins or protein spots are searched which show a difference in their expression levels between for example patients and healthy controls. Most diagnostic markers which are already used in the clinical routine show disease-related changes in a 0.5 to 2-fold range of its biological variation. Thus, to be able to detect new disease-related expression changes (Δ) in a proteome it is necessary to have an analytical detection system which is accurate enough to measure its standard deviation in healthy subjects (σ). The σ value is expected to be in patients at least as high as in healthy subjects. As shown above, however, every protein species of the platelet proteome has its specific variation between healthy subjects.

The significance of an expression level change depends on the effect size (Δ/σ) as well as on the sample size (n per group). Thus, the sensitivity of a study to detect low effect sizes increases with the number of patients and controls. Based on a number of 500 spots, which is in the range of repeatedly found spots in our gels, we calculated the statistical power as a function of the sample size for varying effect sizes (figure 5). Due to the problem of multiple testing, the statistical power depends on the number of spots. The expression level of any spot is certainly not completely independent of the expression levels of all other spots. Therefore we corrected for multiplicity using the False Discovery Rate method according to Benjamini and Hochberg [24]. The results show that an inclusion of 20 patients and 20 controls in a study with platelet protein extracts would allow the detection of an effect size of 1.5 with a power of about 90%. However, an effect size of 1 would have only a power below 40% which is certainly too low. To achieve a sufficient power (80%) for an effect size of 1 a sample size of at least 34 per group would be needed.

Discussion:

In this study, we assessed the technical and the total variation observed for a 2D-DIGE analysis of platelet protein extracts. The technical variation based on a six-fold analysis of the same platelet protein sample was found to be $\pm 5\%$ in SA for a typical spot. The total variation between individual samples was found to be $\pm 16\%$ in SA for a typical spot, based on an analysis of 20 individuals.

For the determination of the total variability, those spots, which were detected and matched in at least 18 of 20 gels (meaning 18 of 20 individuals), were used. This set of spots was also used to determine the technical variability. The constraint to use only those spots, that have been detected in 90% of all gels, results in an efficient removal of those spot features which are just noise, but unfortunately also excludes very faint spots which are not always correctly detected. The relation of the number of gels a spot has to be detected in, and the total number of spots matching this constraint, is already described [25-27]. Of course, the ability of the software to detect spots with constant and good quality, even if total intensity and background levels change (between gels or within gels) and the ability to perform correct automated matching in the presence of spot pattern artefacts and areas of bad resolution, heavily influences the results. A similar number of spots which can be reproducibly detected and correctly matched across many gels, even if the single gels contain 2000-3000 spots, was reported in various other studies [28-30].

Several studies have investigated the technical variation experienced when using two-dimensional gel electrophoresis to analyse the proteome of murine primary cell cultures [31], plant tissues [32, 33], murine brain tissue [30] and porcine blood mononuclear cells [34]. All these studies used conventional methods to stain the proteins, but it is not always clear if any normalisation of the spot volume data was carried out or not. The findings for technical variation were average coefficients of variance (CV) of 16-26% in Coomassie stained gels [32-34], 23% in SyproRuby stained gels [31] and 15% in silver stained gels [30]. As the SA-range calculated for technical variation of 2D DIGE and platelet proteins ($1\pm 5\%$) is the median $\log_2(\text{SA}) \pm \text{median } \sigma$, that was back-transformed from logarithmic space, a CV of about 5% can be calculated, to obtain a result comparable with the studies mentioned before. Obviously, the usage of the internal standard dramatically reduces the technical variability. This effect has previously been demonstrated for selected spots

in a 2D map where similar values for CV were obtained when the internal standard was used [35]. Alternative normalisation methods which could further decrease the technical variability have been presented [36, 37]. These methods would be especially beneficial for small volume spots, but a large part of these spots has been removed because of the spot selection criteria described above, so the alternative normalisation methods were not applied to our data.

Several studies which assessed the technical variation, also determined the total (inter-individual) variation experienced when comparing samples taken from different individuals: The total variation using conventional staining methods was found to be in murine primary cell lines CV=40-46% (tech: 23%) [31], plant tissues CV=24% (tech: 16%) and 40-60% (tech: 26%) [32] and [33], porcine mononuclear cells CV=34% (tech 15%) [34]. These findings indicate a significant increase in spot volume variation when samples, which originate from different organisms are analysed. Our results for the total variation of SA $1\pm 16\%$ would give a CV of 16%, which is significantly lower than the values found by the studies mentioned above, but the increase in CV from technical to total variation is similar. A recent study where the liver proteome has been analysed using 2D-DIGE, has also analysed the total variation and found a median CV of SA values of 19%, which perfectly matches our findings. Our data also show a by far broader distribution in total variation than observed for technical variation. Since the technical and the total variation do not correlate, nor do the total correlation and pI, MW and spot volume (protein abundance), it has to be assumed, that some protein species show rather high differences in protein abundance levels in different individuals. This effect can also be observed in human plasma where the CV of the protein concentration of selected proteins ranges from 14% to 93% within a group of different individuals [38]. It was not in the scope of this study to analyse the variation experienced between samples that have been consecutively taken from the same individual, so this type of variation contributes in an unknown amount to the total variation. Due to standardised sampling conditions and sample preparation procedures, this type of variation is considered to be low and of constant height. This assumption is supported by finding of Nelsestuen *et al.* [39], who compared plasma protein profiles obtained by MALDI mass spectrometry from samples of different individuals, which were taken at certain time points. They describe that the protein expression levels of each individual

change very little over time and the observed differences between individuals are significantly higher at any time point. We also detected spots which show only a slightly higher total variation than technical variation. The proteins in these spots would be good candidates to be used for normalisation or load control in methods which don't allow the use of internal standards. There were 18 spots out of this set selected and identified by mass spectrometry at the state-of-the-art level. Table 2 lists the complete set of data used for the identification of the selected protein spots.

The mass spectrometric analysis of spot 2047 indicated the simultaneous presence of 2 proteins with high homology (81% identity), which obviously have migrated to the same position on the 2D gel, which is observed at times [40]. For spot 1518, the obtained mass spectrometric data were not sufficient to either confirm or deny the presence of other highly homologous (identical sequences except N-terminus) isoforms of protein disulfide-isomerase. The results for spot 1652 were not satisfying, because the obtainable data were not sufficient to distinguish between the two isoforms Nucleoside diphosphate kinase A and B because of their high sequence homology. Spots 1671 and 1672 have both been identified as Cyclophilin A and spots 1177 and 1200 as Glycerinaldehyde-3-phosphate dehydrogenase (GAPDH), suggesting the presence of different posttranslational modifications for these proteins. Isoforms of all suggested proteins could be present and show considerable variation between individuals, so it is advisable to validate the stability in protein expression levels of all candidate proteins, if different tissues or quantification methods are used. The RNA transcripts of some of the suggested proteins are already used to do normalisation in real-time PCR, especially the GAPDH transcript is frequently used, but its expression level shows significant differences across different tissues, but is relatively stable when comparing the same tissue of different individuals [41]. It is suggested rather to use a panel of transcripts for normalisation in real-time PCR than one alone [42]. This suggestion obviously also applies to western blotting, as it has been demonstrated, that the use of the proteins GAPDH, β -actin and β -tubulin alone is not suitable to perform normalisation or load control due to drastic differences in the expression level between different tissues or disease states [43, 44].

Expression changes observed in clinical studies, are mainly in the range of 1.5 to 2-fold. Higher expression changes are seldom reached. With a technical variation of

5% 2D-DIGE is able to detect clinically relevant expression changes. However, based on our results for the total variation experienced between individuals, we suggest not to focus on minimum expression changes, but to consider the use of the effect size to evaluate the statistical significance for each spot at the given sample number.

References:

1. LaBaer, J., So, you want to look for biomarkers (introduction to the special biomarkers issue). *J Proteome Res*, 2005. **4**(4): p. 1053-9.
2. Jurk, K. and B.E. Kehrel, *Platelets: physiology and biochemistry*. *Semin Thromb Hemost*, 2005. **31**(4): p. 381-92.
3. Richardson, J.L., et al., *Mechanisms of organelle transport and capture along proplatelets during platelet production*. *Blood*, 2005. **106**(13): p. 4066-75.
4. Gnatenko, D.V., et al., *Transcript profiling of human platelets using microarray and serial analysis of gene expression*. *Blood*, 2003. **101**(6): p. 2285-93.
5. Thom, T., et al., *Heart disease and stroke statistics--2006 update: a report from the American Heart Association Statistics Committee and Stroke Statistics Subcommittee*. *Circulation*, 2006. **113**(6): p. e85-151.
6. Wilkerson, W.R. and D.C. Sane, *Aging and thrombosis*. *Semin Thromb Hemost*, 2002. **28**(6): p. 555-68.
7. Unlu, M., M.E. Morgan, and J.S. Minden, *Difference gel electrophoresis: a single gel method for detecting changes in protein extracts*. *Electrophoresis*, 1997. **18**(11): p. 2071-7.
8. Marcus, K., et al., *Identification of platelet proteins separated by two-dimensional gel electrophoresis and analyzed by matrix assisted laser desorption/ionization-time of flight-mass spectrometry and detection of tyrosine-phosphorylated proteins*. *Electrophoresis*, 2000. **21**(13): p. 2622-36.
9. Garcia, A., et al., *Extensive analysis of the human platelet proteome by two-dimensional gel electrophoresis and mass spectrometry*. *Proteomics*, 2004. **4**(3): p. 656-68.
10. Coppinger, J.A., et al., *Characterization of the proteins released from activated platelets leads to localization of novel platelet proteins in human atherosclerotic lesions*. *Blood*, 2004. **103**(6): p. 2096-104.
11. Maguire, P.B., et al., *Identification of the phosphotyrosine proteome from thrombin activated platelets*. *Proteomics*, 2002. **2**(6): p. 642-8.
12. Marcus, K., J. Moebius, and H.E. Meyer, *Differential analysis of phosphorylated proteins in resting and thrombin-stimulated human platelets*. *Anal Bioanal Chem*, 2003. **376**(7): p. 973-93.
13. Zellner, M., et al., *Quantitative validation of different protein precipitation methods in proteome analysis of blood platelets*. *Electrophoresis*, 2005. **26**(12): p. 2481-9.

14. Shevchenko, A., et al., *Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels*. Anal Chem, 1996. **68**(5): p. 850-8.
15. Jensen, O.N., et al., *Sample preparation methods for mass spectrometric peptide mapping directly from 2-DE gels*. Methods Mol Biol, 1999. **112**: p. 513-30.
16. Belgacem, O., et al., *Dissociation of biomolecules using a ultraviolet matrix-assisted laser desorption/ionisation time-of-flight/curved field reflectron tandem mass spectrometer equipped with a differential-pumped collision cell*. Rapid Commun Mass Spectrom, 2006. **20**(11): p. 1653-60.
17. Vorm, O., P. Roepstorff, and M. Mann, *Improved Resolution and Very High Sensitivity in MALDI TOF of Matrix Surfaces Made by Fast Evaporation*. Analytical Chemistry, 1994. **66**(19): p. 3281-7.
18. Mattow, J., et al., *Protein identification and tracking in two-dimensional electrophoretic gels by minimal protein identifiers*. Proteomics, 2004. **4**(10): p. 2927-41.
19. Smirnov, I.P., et al., *Suppression of alpha-cyano-4-hydroxycinnamic acid matrix clusters and reduction of chemical noise in MALDI-TOF mass spectrometry*. Anal Chem, 2004. **76**(10): p. 2958-65.
20. Zhang, W. and B.T. Chait, *ProFound: an expert system for protein identification using mass spectrometric peptide mapping information*. Anal Chem, 2000. **72**(11): p. 2482-9.
21. Perkins, D.N., et al., *Probability-based protein identification by searching sequence databases using mass spectrometry data*. Electrophoresis, 1999. **20**(18): p. 3551-67.
22. Bairoch, A. and R. Apweiler, *The SWISS-PROT protein sequence database: its relevance to human molecular medical research*. J Mol Med, 1997. **75**(5): p. 312-6.
23. Savitzky, A. and M. Golay, *Smoothing and Differentiation of Data by Simplified Least Squares Procedures*. Analytical Chemistry, 1964. **36**(8): p. 1627-1639.
24. Benjamini, Y. and Y. Hochberg, *Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing*. Journal of the Royal Statistical Society B, 1995. **57**(1): p. 289-300.
25. Corzett, T.H., et al., *Statistical analysis of the experimental variation in the proteomic characterization of human plasma by two-dimensional difference gel electrophoresis*. J Proteome Res, 2006. **5**(10): p. 2611-9.
26. Fodor, I.K., et al., *Statistical challenges in the analysis of two-dimensional difference gel electrophoresis experiments using DeCyder*. Bioinformatics, 2005. **21**(19): p. 3733-40.

27. Voss, T. and P. Haberl, *Observations on the reproducibility and matching efficiency of two-dimensional electrophoresis gels: consequences for comprehensive data analysis*. *Electrophoresis*, 2000. **21**(16): p. 3345-50.
28. Zhang, X., et al., *Proteomic analysis of individual variation in normal livers of human beings using difference gel electrophoresis*. *Proteomics*, 2006. **6**(19): p. 5260-8.
29. Terry, D.E. and D.M. Desiderio, *Between-gel reproducibility of the human cerebrospinal fluid proteome*. *Proteomics*, 2003. **3**(10): p. 1962-79.
30. Challapalli, K.K., et al., *High reproducibility of large-gel two-dimensional electrophoresis*. *Electrophoresis*, 2004. **25**(17): p. 3040-7.
31. Molloy, M.P., et al., *Overcoming technical variation and biological variation in quantitative proteomics*. *Proteomics*, 2003. **3**(10): p. 1912-9.
32. Asirvatham, V.S., B.S. Watson, and L.W. Sumner, *Analytical and biological variances associated with proteomic studies of *Medicago truncatula* by two-dimensional polyacrylamide gel electrophoresis*. *Proteomics*, 2002. **2**(8): p. 960-8.
33. Jorge, I., et al., *The holm oak leaf proteome: analytical and biological variability in the protein expression level assessed by 2-DE and protein identification tandem mass spectrometry de novo sequencing and sequence similarity searching*. *Proteomics*, 2005. **5**(1): p. 222-34.
34. Ramirez-Boo, M., et al., *Analysis of porcine peripheral blood mononuclear cells proteome by 2-DE and MS: Analytical and biological variability in the protein expression level and protein identification*. *Proteomics*, 2006. **6 Suppl 1**: p. S215-25.
35. Alban, A., et al., *A novel experimental design for comparative two-dimensional gel analysis: two-dimensional difference gel electrophoresis incorporating a pooled internal standard*. *Proteomics*, 2003. **3**(1): p. 36-44.
36. Karp, N.A., D.P. Kreil, and K.S. Lilley, *Determining a significant change in protein expression with DeCyder during a pair-wise comparison using two-dimensional difference gel electrophoresis*. *Proteomics*, 2004. **4**(5): p. 1421-32.
37. Kreil, D.P., N.A. Karp, and K.S. Lilley, *DNA microarray normalization methods can remove bias from differential protein expression analysis of 2D difference gel electrophoresis results*. *Bioinformatics*, 2004. **20**(13): p. 2026-34.
38. Anderson, N.L. and N.G. Anderson, *The human plasma proteome: history, character, and diagnostic prospects*. *Mol Cell Proteomics*, 2002. **1**(11): p. 845-67.
39. Nelsestuen, G.L., et al., *Plasma protein profiling: unique and stable features of individuals*. *Proteomics*, 2005. **5**(15): p. 4012-24.
40. Gygi, S.P., et al., *Evaluation of two-dimensional gel electrophoresis-based proteome analysis technology*. *Proc Natl Acad Sci U S A*, 2000. **97**(17): p. 9390-5.

41. Barber, R.D., et al., *GAPDH as a housekeeping gene: analysis of GAPDH mRNA expression in a panel of 72 human tissues*. *Physiol Genomics*, 2005. **21**(3): p. 389-95.
42. Tricarico, C., et al., *Quantitative real-time reverse transcription polymerase chain reaction: normalization to rRNA or single housekeeping genes is inappropriate for human tissue biopsies*. *Anal Biochem*, 2002. **309**(2): p. 293-300.
43. Dittmer, A. and J. Dittmer, *Beta-actin is not a reliable loading control in Western blot analysis*. *Electrophoresis*, 2006. **27**(14): p. 2844-5.
44. Ferguson, R.E., et al., *Housekeeping proteins: a preliminary study illustrating some limitations as useful references in protein expression studies*. *Proteomics*, 2005. **5**(2): p. 566-71.

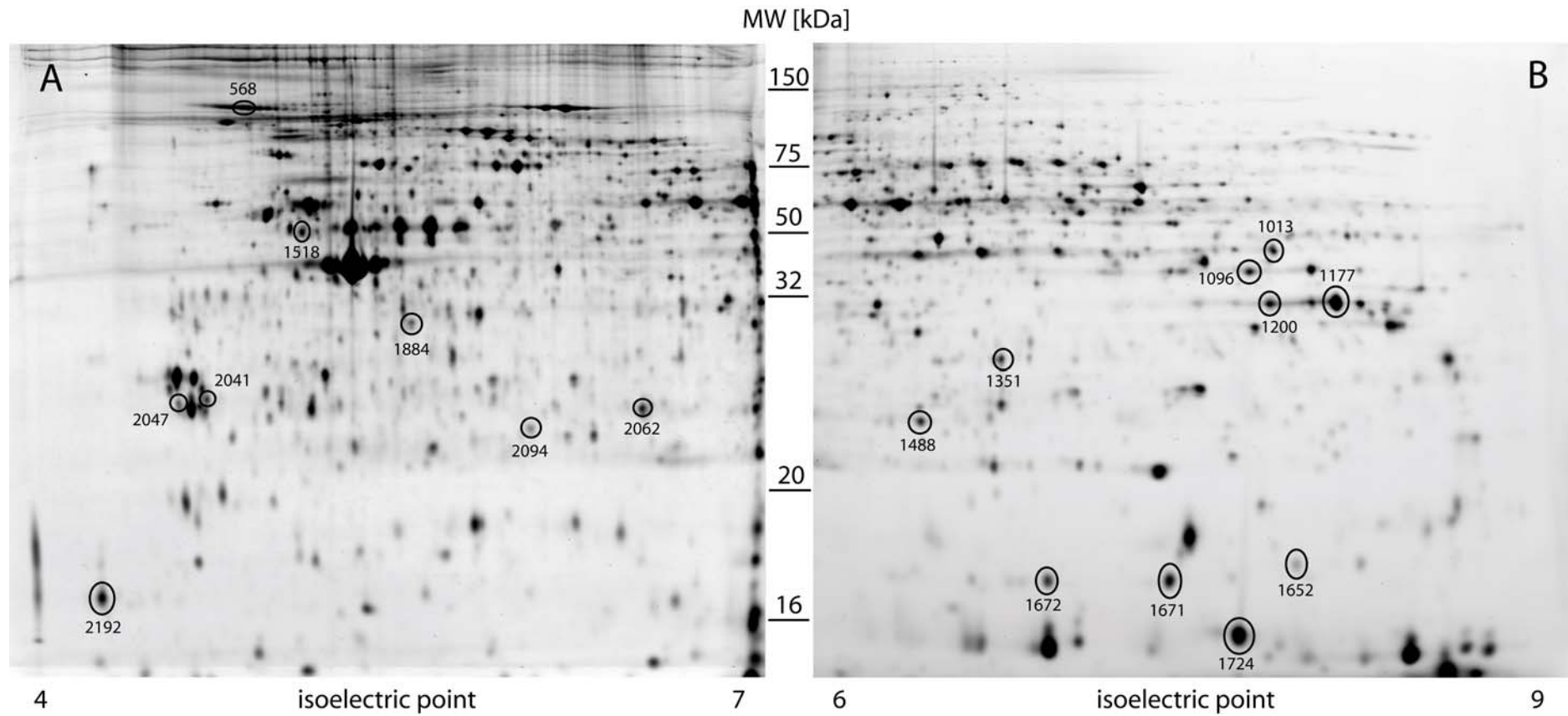


Figure 1

Typical 2D images for pI ranges 4-7 and 6-9 (Channel: Cy2, internal standard, 33 μ g protein). Circles indicate spots which were found to show low total variability (standard deviation of $\log_2(\text{SA}) < 0.125$) and have been identified by mass spectrometry.

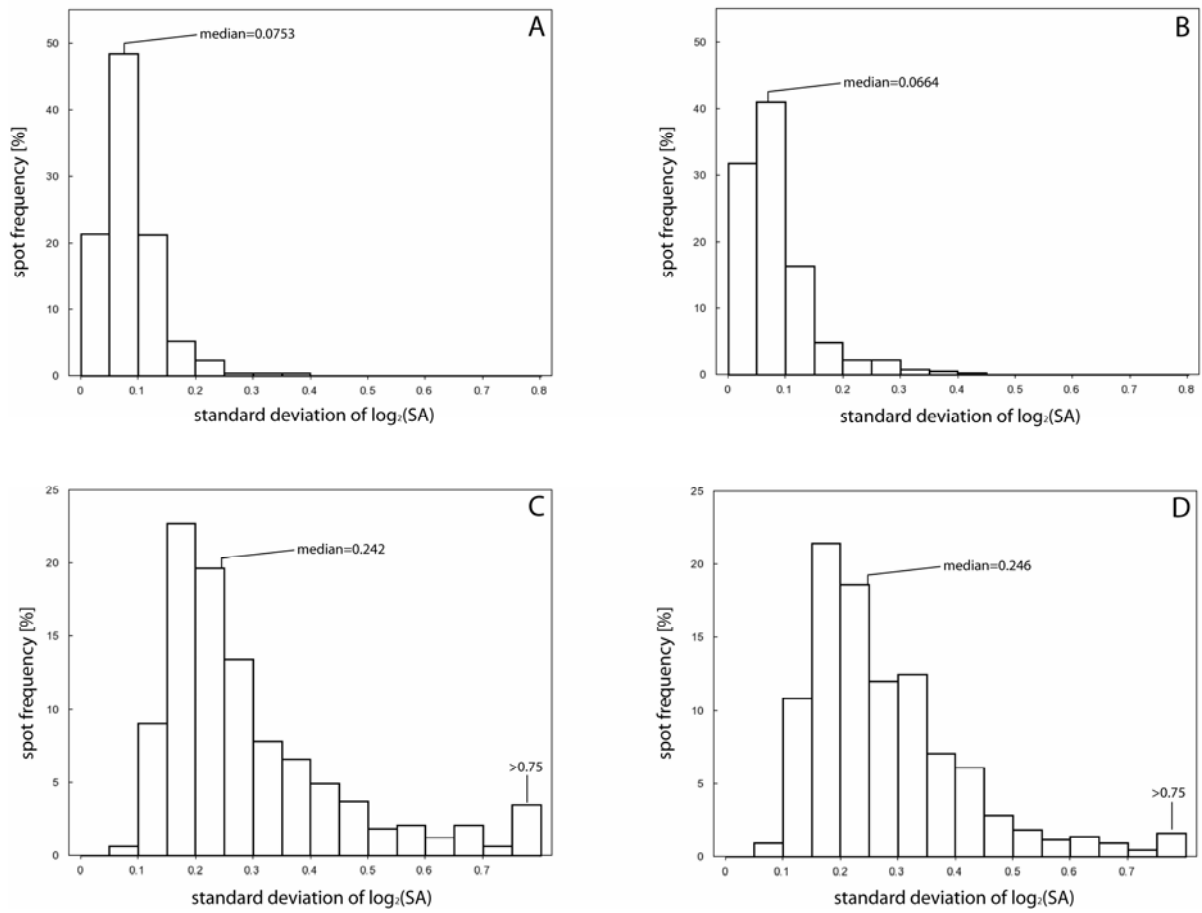


Figure 2.

Frequency distribution of the standard deviation of $\log_2(\text{SA})$ visualising the technical variability for pl-ranges 4-7(A) and 6-9(B) and the total variability for pl-ranges 4-7(C) and 6-9(D)

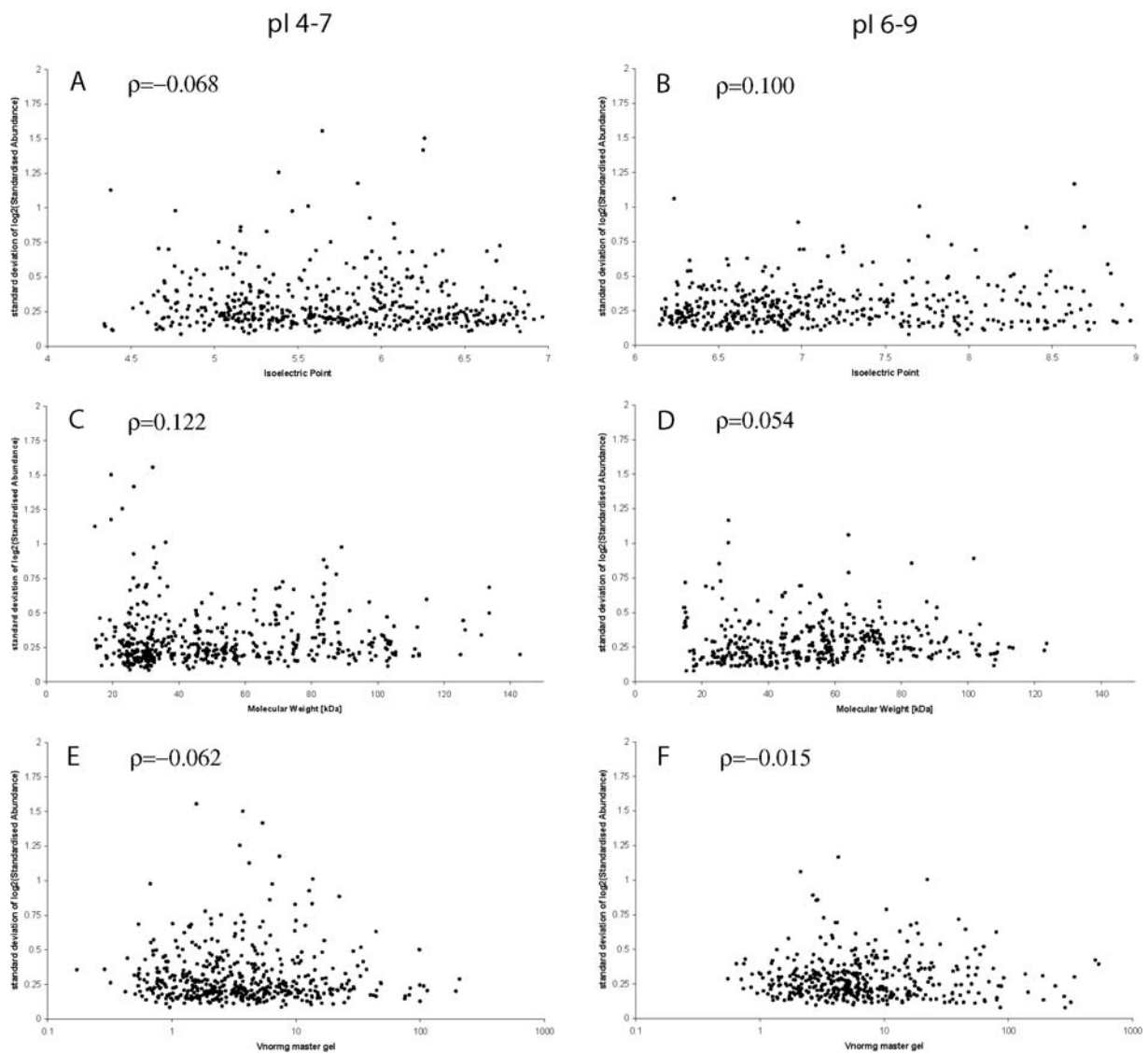


Figure 3

Scatter plots to visualise the independence of the standard deviation from pI (A,B), molecular weight (C,D) and spot volume (E,F). Spearman rank correlation coefficients (ρ) have been calculated to confirm the independence of the standard deviation.

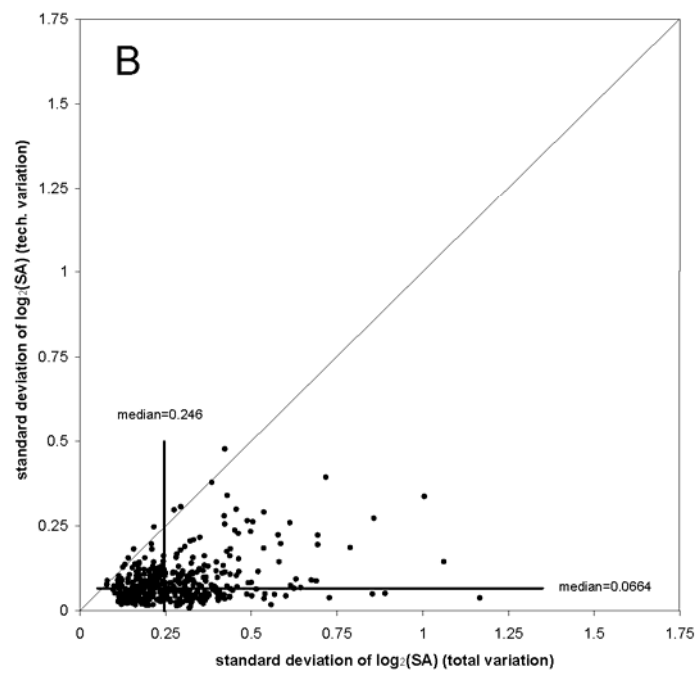
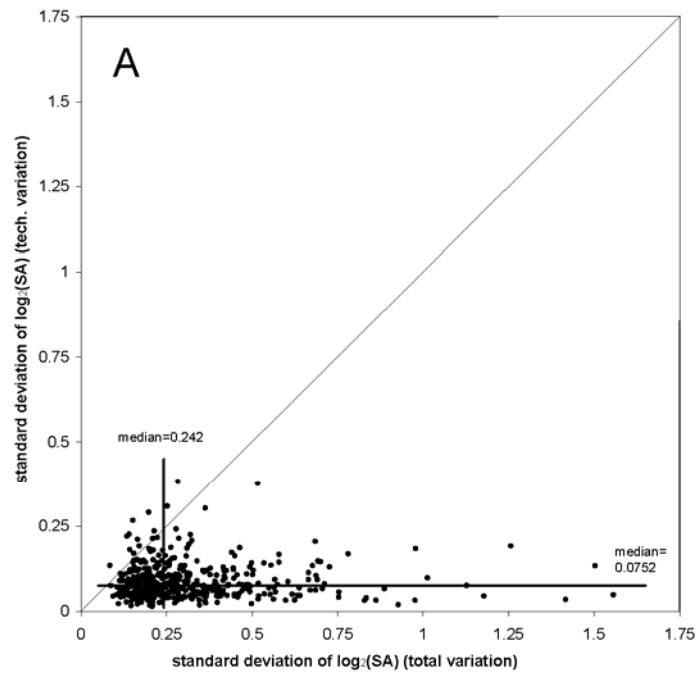


Figure 4. Correlation plots of technical and total variation for pI ranges 4-7 (A) and 6-9 (B). The standard deviation determined for total the variation has been plotted against the standard deviation for technical variation. Medians are drawn as bold lines. Spots below the 45° line show higher total variance than technical variance, which is true for the vast majority of all spots.

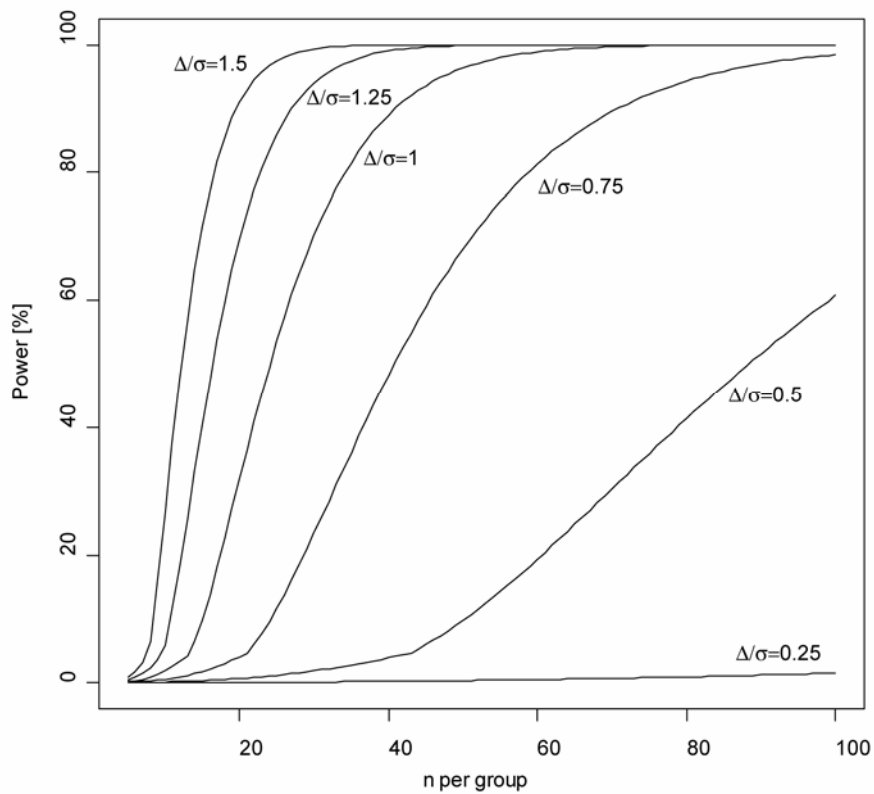


Figure 5

Dependence of the power for a 2-sided unpaired t-test controlling a false discovery rate of 0.05 on the sample size per group for different effect sizes (Δ/σ). 20 of 500 spots were assumed to be changes in expression.

Tables:

Sample	Age	sex	Dye 4-7	Dye 6-9
1	100	f	Cy5	Cy5
2	89	f	Cy5	Cy5
3	82	f	Cy3	Cy5
4	87	f	Cy5	Cy3
5	72	m	Cy3	Cy3
6	88	f	Cy5	Cy5
7	82	f	Cy5	Cy5
8	84	f	Cy3	Cy5
9	82	f	Cy3	Cy3
10	87	f	Cy3	Cy3
11	66	f	Cy3	Cy3
12	71	f	Cy5	Cy5
13	73	f	Cy5	Cy3
14	93	m	Cy5	Cy5
15	59	m	Cy5	Cy5
16	56	m	Cy3	Cy3
17	62	m	Cy5	Cy5
18	75	f	Cy5	Cy5
19	73	f	Cy3	Cy3
20	64	m	Cy3	Cy3

Table 1
List of individuals and labelling strategy

Spot id	pI/MW on gel ⁵	pI/MW calculated ⁶	avg V_{norm_g}	total: avg $\log_2(SA)$	total: SD of $\log_2(SA)$	tech: SD of $\log_2(SA)$	Protein Name	Swissprot Acc. Nr.	Sequence Coverage PMF	Matched peptides PMF	MOWSE Score PMF ³	Peptides MS/MS	MOWSE Score MS/MS ⁴
568	4.9 / 103	5.2 / 110	21.10	-0.0828	0.1152	0.0421	Integrin α -IIb	P08514	27%	22/26	241	4	211
1518	5.2 / 44		14.91	0.0616	0.1057	0.0496	Protein disulfide-isomerase related prot. 5 ¹	nr: 1710248 ²	36%	13/20	134	2	66
1884	5.6 / 30	5.6 / 37	4.68	0.0205	0.1041	0.0828	Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta 1 ¹	P62873	14%	5/10	63	3	62
2041	4.8 / 25	4.8 / 28	8.99	-0.0106	0.0872	0.0753	14-3-3 protein gamma	P61981	32%	12/15	135	2	52
2047	4.7 / 25	4.7 / 28	6.44	-0.0275	0.1233	0.0900	14-3-3 protein theta	P27348	40 %	15/30	151		
		4.7 / 28					14-3-3 protein zeta/delta	P63104	31%	10/30	84		
2062	6.5 / 25	6.0 / 25	15.77	0.0829	0.1092	0.0226	Peroxiredoxin-6	P30041	13%	4/5	69	1	61
2094	6.1 / 24	5.9 / 25	4.50	0.0689	0.1235	0.0588	Growth factor receptor-bound protein 2	P62993	17%	5/6	80	3	68
2192	4.4 / 17	4.6 / 17	36.92	-0.0251	0.1186	0.1038	Myosin light polypeptide 6	P60660	52%	11/30	115	1	48
1013	8.1 / 46	8.3 / 47	24.96	0.1042	0.1117	0.0270	Isocitrate Dehydrogenase	P48735	28%	11/15	142	3	102
1096	8.0 / 41	8.4 / 39	193.39	0.0234	0.1166	0.0474	Fructose-bisphosphate aldolase A	P04075	44%	15/20	221	2	156
1177	8.4 / 35	8.6 / 36	26.20	-0.0237	0.1171	0.0309	Glyceraldehyde-3-phosphate dehydrogenase	P04406	59%	14/30	168	3	152
1200	8.1 / 35	8.6 / 36	38.05	-0.0692	0.1228	0.0898	Glyceraldehyde-3-phosphate dehydrogenase	P04406	24%	8/11	118	1	70
1351	6.9 / 30	6.8 / 34	20.73	0.0911	0.1163	0.0411	Actin-related protein 2/3 complex subunit 2	O15144	19%	9/18	94	2	73
1488	6.6 / 26	6.5 / 27	22.89	0.0508	0.1219	0.0513	Triosephosphate-Isomerase	P60174	71%	15/28	162	2	227
1652	8.2 / 18	8.5 / 17	9.78	0.0561	0.1152	0.0664	Nucleoside diphosphate kinase A or B	P15531 or P22392	19%	3/3	59	2	72
1671	7.6 / 18	7.8 / 18	60.63	-0.0118	0.0794	0.0890	Cyclophilin A	P62937	42%	6/12	95		
1672	7.1 / 18	7.8 / 18	34.98	-0.0741	0.1110	0.0784	Cyclophilin A	P62937	38%	7/11	99	3	51
1724	7.9 / 15	8.5 / 15	216.09	0.0080	0.0785	0.0769	Profilin-1	P07737	58%	8/15	135	2	69

Table 2

Protein identification results for spots with standard deviations of $\log_2(SA) < 0.125$, see figure 1 for location on 2D-gels.

All spots were detected in 20/20 gels except 1884 (4-7) and 1652 (6-9) which were found in 19/20 gels

- (1) PMF results suggest the co-migration of highly homologous isoforms (>95% identity), but neither presence nor absence could be confirmed
- (2) Accession-nr. from NCBI nr database
- (3) Significance threshold for MASCOT MOWSE score (PMF): Searching SWISSPROT database: 54, Searching NCBI nr: 64
- (4) Significance threshold for MASCOT MOWSE score (MS/MS): Searching SWISSPROT database: 27, Searching NCBI nr: 36
- (5) Values determined after calibration using theoretical pI/MW of already known landmark proteins (data not shown)
- (6) Calculated using the expasy pi_tool (http://www.expasy.ch/tools/pi_tool.html) for the whole sequence (omitting signal peptides if annotated).

Conclusion

The biomarker-discovery workflow as it was used in this PhD-thesis, consists of initial sampling and sample preparation, followed by protein separation using two-dimensional gel electrophoresis with subsequent quantification based on fluorescence dye labelling. The quantification data are analysed to find protein spots, which show different expression levels when comparing two groups of samples (e.g. treated and untreated cells). These spots are selected and the proteins were identified by mass spectrometry.

This workflow could be successfully adopted to detect and identify proteins whose expression levels change during glutamine starvation in the monocytic cell line U937. Several spots were detected, which clearly showed a change in expression level. In some cases, these changes were even by factor of four. Changes of this degree have to be considered as rather high compared to the results of other biomarker searches, but glutamine starvation is a model to simulate the presence of a critical disease - so these results meet the expectations. Nearly all spots of interest, except the faintest ones, could be successfully identified by peptide mass fingerprinting using a matrix-assisted laser desorption/ionisation reflectron time of flight instrument and these identification results were further confirmed by MS/MS experiments, delivering at least one sequence tag.

The results of this study gave valuable experiences, which could be applied to samples from primary patient material. Studies based on these samples always have to deal with limited material and complex sample composition, and therefore the next step was the optimisation and validation of the sample preparation procedure from human thrombocytes. Thrombocytes cannot be concentrated by centrifugation, because this can cause activation of uncertain degree, which would introduce changes in their proteome. So, to obtain a sample with a protein concentration suitable for two-dimensional gel electrophoresis, the platelet proteins had to be precipitated and subsequently resolubilised in sample buffer. The use of both precipitation agents, ethanol and trichloroacetic acid gave samples which were compatible to two-dimensional electrophoresis, but trichloroacetic acid proved to be

superior to ethanol, because the samples don't have to be dialysed prior to electrophoresis.

This preliminary work has made it possible to analyse the variability of the thrombocyte protein pattern by two-dimensional differential gel electrophoresis (2D-DIGE). The knowledge of the extent of the technical as well as the total variability is of great importance for planning a study to search for biomarkers. The technical variability defines the limit for minimally detectable expression changes. This type of variability was determined to be $\pm 5\%$ for a typical spot, which shows the superiority of 2D-DIGE and of the application of an internal standard over conventional staining methods. The total variability defines the number of samples that are needed to detect given expression changes with statistical significance. It was determined within a group of 20 healthy volunteers (aged from 56-100 years) and found to be $\pm 16\%$ for a median spot. A broad distribution of the extent of the total variability has been found, indicating that some protein species shows rather high variability within the considered group. Additionally, a group of protein spots could be detected, which showed a hardly higher variability than is was determined for the technical variability. These spots were excised and their proteins were identified by mass spectrometry via peptide mass fingerprinting and subsequent high energy CID MS/MS experiments to confirm the identification results. The proteins out of this set can be used to do load control or normalisation in experiments that don't apply internal standards. With the findings for the total variation, it was possible to provide an estimation of the correlation of the statistical power and the sample size at given effect sizes.

Taking into account these results we are now able to define the technical procedure to perform a biomarker search in blood cells without further experiments.

Lebenslauf

Mag. Wolfgang Winkler
Jagdschlossgasse 54/9/7
1130 Wien

Geboren am: 14.06.1977, Wien

- 12/2003 – 9/2006 Dissertation in den Arbeitsgruppen von
Prof. Dr. Günter Allmaier
Institut für Chemische Technologien und Analytik
Technische Universität Wien
und
Prof. Dr. Rudolf Oehler
Forschungslaboratorien der Universitätsklinik für Chirurgie
Medizinische Universität Wien
Titel: *Establishment of a proteomic workflow for biomarker research in
blood cell samples*
- 12/2002 – 11/2003 Diplomarbeit in der Arbeitsgruppe von
Prof. Dr. Rudolf Oehler
Forschungslaboratorien der Universitätsklinik für Chirurgie
Medizinische Universität Wien
Titel: *Proteom-Analyse an Thrombocyten alter und junger Menschen*
- 10/1997 – 04/2004 Studium Chemie
Schwerpunkt Biochemie, Analytische und Organische Chemie
Universität Wien
- 01/1997 – 08/1997 Präsenzdienst
- 09/1991 – 06/1996 HBLVA f. Chemische Industrie, Ausbildungszweig Biochemie
Rosensteingasse 79, 1170 Wien

Publikationen

1. Quantitative validation of different protein precipitation methods in proteome analysis of blood platelets
Zellner M, Winkler W, Hayden H, Diestinger M, Eliassen M, Gesslbauer B, Miller I, Chang M, Kungl A, Roth E, Oehler R
Electrophoresis. 2005 (12), 2481-2489
2. Adaptive cellular mechanisms in response to glutamine-starvation
Eliassen M, Winkler W, Jordan V, Pokar M, Marchetti M, Roth E, Allmaier G, Oehler R
Frontiers in Bioscience. 2006 (11), 3199-3211
3. Biological variation of the platelet proteome in the elderly population and its implication for biomarker research (In preparation)
Winkler W, Zellner M, Diestinger M, Marchetti M, Allmaier G, Oehler R

Vorträge

1. Biological variation of the platelet proteome in the elderly population and its implication for biomarker research
Winkler W, Zellner M, Diestinger M, Rappold E, Goll A, Zehetmayer S, Bauer P, Marchetti M, Allmaier G, Oehler R
4. Österreichisches Proteomforschungssymposium; Vienna, Austria: 20. – 21. Sept. 2006

Poster

1. Platelet Proteomics: A tool to analyse age-related changes in the protein expression profile
Winkler W, Zellner M, Roth E, Oehler R
21. Kongress der AKE, DGEM und GESKES; Linz, 12. - 14. Jun. 2003
2. Effect of different protein precipitation methods on results in proteome analysis of human blood platelets
Winkler W, Zellner M, Miller I, Chang M, Oehler R
Proteomic Forum; München, 14. - 17. Sept. 2003
3. Platelet Proteomics: A tool to analyse age-related changes in the protein expression profile
Zellner M, Winkler W, Miller I, Chang M, Grillari J, Staes A, Hugelier K, Vandekerckhove J, Gevaert K, Oehler R
6th Siena-meeting: From Proteome to Genome; Siena, 30. Aug. – 3. Sept. 2004
4. Glutamine starvation induces specific adaptive response in monocytic cells
Eliassen M, Winkler W, Gerner C, Zehl M, Marchetti M, Allmaier G, Oehler R
2nd International Symposium of the Austrian Proteomics Platform: Seefeld/Tirol, 24. - 27. Jän. 2005
5. Fever-range hyperthermia induces specific proteome changes in human lymphocytes
Radwan M, Eliassen M, Winkler W, Seipelt J, Sommergruber W, Zehl M, Marchetti M, Allmaier G, Oehler R
2nd International Symposium of the Austrian Proteomics Platform; Seefeld/Tirol, 24. - 27. Jän. 2005
6. Protein profiling of human blood platelets from Alzheimer's disease patients
Diestinger M, Zellner M, Rappold E, Winkler W, Gesslbauer B, Garo F, Oehler R
HUPO 4th annual World Congress; München, 28. Aug. – 1. Sept. 2005
Proceedings: Molecular & Cellular Proteomics, 4 (2005), 8; Supplement S146
7. Protein profiling of human blood platelets from young and aged individuals
Winkler W, Zellner M, Rappold E, Garo F, Miller I, Chang M, Grillari J, Staes A, Hugelier K, Vandekerckhove J, Gevaert K, Oehler R
HUPO 4th annual World Congress; München, 28. Aug. – 1. Sept. 2005
Proceedings: Molecular & Cellular Proteomics, 4 (2005), 8; Supplement S167
8. Protein profiling of human blood platelets from idiopathic Parkinson's disease patients
Zellner M, Winkler W, Kotzailias N, Gesslbauer B, Sycha T, Diestinger M, Oehler R
HUPO 4th annual World Congress; München, 28. Aug. – 1. Sept. 2005
Proceedings: Molecular & Cellular Proteomics, 4 (2005), 8; Supplement S168